

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

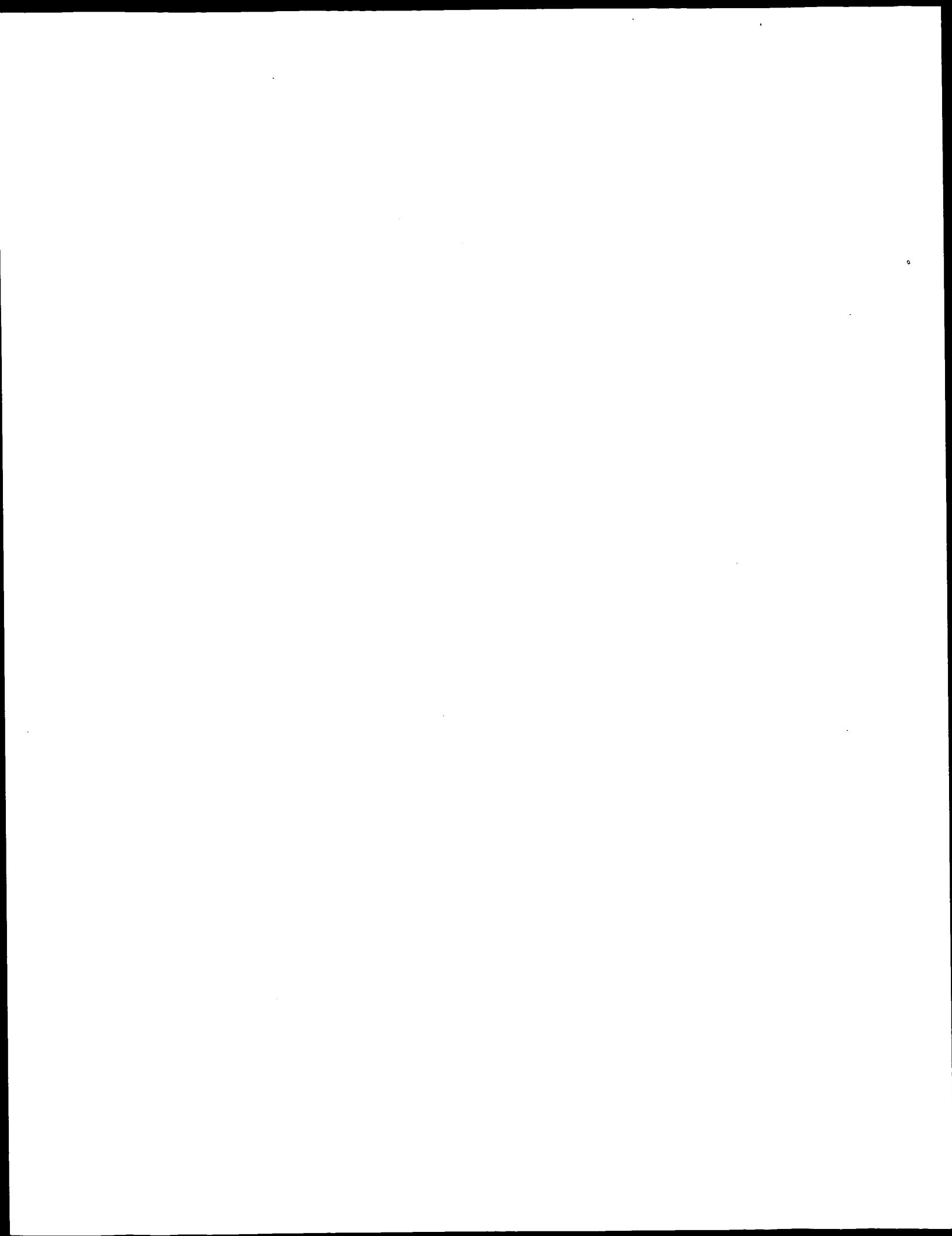
Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 04-16-03
Searcher: Beverly C 4944
Terminal time: 23
Elapsed time: _____
CPU time: _____
Total time: 23
Number of Searches: _____
Number of Databases: 1

Search Site	Vendors
<input type="checkbox"/> STIC	<input type="checkbox"/> IG Suite
<input type="checkbox"/> CM-1	<input type="checkbox"/> STN
<input type="checkbox"/> Pre-S	<input type="checkbox"/> Dialog
Type of Search	
<input type="checkbox"/> N.A. Sequence	<input type="checkbox"/> APS
<input type="checkbox"/> A.A. Sequence	<input type="checkbox"/> Geninfo
<input type="checkbox"/> Structure	<input type="checkbox"/> SDC
<input type="checkbox"/> Bibliographic	<input checked="" type="checkbox"/> DARC/Questel
	<input checked="" type="checkbox"/> Other C G W



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Run on: April 15, 2003, 13:06:01 ; Search time 3378 Seconds
(without alignments)
9175.396 Million cell updates/sec

Title: US-09-905-173-23
Sequence: 1 atgatacccaagggattaa.....cccttcaagctctttaa 1065

Scoring table: IDENTITY_NUC
Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Result No.	Score	Query Match	Length	DB ID	Description
1	1065	10.0	1065	6 AR043346	AR043346 Sequence
2	1065	10.0	1065	6 AR078122	AR078122 Sequence
3	1063.4	9.9	8 12497	1 AR000771	AE000771 Aquifex a
4	145.4	13.7	12075	1 AR013160	AE013160 Thermococcus
5	103.8	9.7	12452	1 AE010265	AE010265 Pyrococcus
6	90.8	8.5	11932	1 AR002508	AE002508 Neisseria
7	90.8	8.5	11932	6 AX044033	AX044033 Sequence
8	89.2	8.4	20334	1 AE001764	AE001764 Thermotoga
9	87.6	8.2	320301	1 NHA627491	AL162757 Neisseria
10	86	8.1	22014	1 AE00963	AE00963 Archaeoglobi
11	80.6	7.6	188050	1 AL646072	AL646072 Ralstonia
12	72.2	6.8	1480	1 AF246314	AF246314 Acidithio
13	68.8	6.5	11265	1 AE012885	AE012885 Chlorobium
14	64.2	6.0	1285	8 AX17255	AY17255 Arabidopsis
15	64.2	6.0	1570	8 AY050832	AY050832 Arabidopsis
16	61	5.7	7218	6 166494	166494 Sequence 14
17	60.2	5.7	17492	1 U67539	U67539 Methanococcus
18	60	5.6	1242	8 NP1278767	AJ278767 Nicotiana
19	56.8	5.6	1556	8 NTH94	Y09204 N tabacum m
20	55.2	5.2	300600	1 AP005369	AP005369 Thermosyn
21	54.4	5.1	13190	1 AB013011	AE013011 Thermomana
22	46.2	4.3	1085	6 AX431883	AX431883 Sequence
23	45	4.2	301150	1 AP003190	AP003190 Clostridi
24	44.2	4.2	11749	1 AE00964	AE00964 Archaeoogl
25	44.2	4.2	25770	1 LJU92974	U92974 Lactococcus
26	43.6	4.1	203335	1 CJI1188X5	AI139078 Campyloba
27	42.2	4.0	1092	6 AR078126	AR078126 Sequence
28	41.6	3.9	119184	1 D90909	D90909 Synechocyst
29	40.8	3.8	10872	1 AE008031	AE008031 Agrobacte
30	40.8	3.8	11490	1 AE009065	AE009065 Agrobacte
31	40.8	3.8	189287	2 AC068362	AC068362 Homo sapi
32	40.6	3.8	200918	9 AC012360	AC012360 Homo sapi
33	40.2	3.8	177585	2 AC094167	AC094167 Rattus no
34	40.2	3.8	226608	2 AC105612	AC105612 Rattus no
35	40	3.8	65785	2 AC10101	AC10101 Rattus no
36	40	3.8	91365	2 AP004776	AP004776
37	40	3.8	181159	2 AL772390	AL772390 Mus muscu
38	40	3.8	158050	8 AF480496	AF480496 Oryza sat
39	39.8	3.7	147565	2 AC112868	AC112868 Rattus no
40	39.8	3.7	176744	2 AC128581	AC128581 Rattus no
41	39.8	3.7	201420	2 AC099110	AC099110 Rattus no
42	39	3.7	11080	1 AE006160	AE006160 Pasteurel
43	39	3.7	215296	2 AC09369	AC09369 Mus muscu
44	38.8	3.6	12257	1 AE009793	AE009793 Pyrobacul
45	38.8	3.6	177014	2 AC022269	AC022269 Homo sapi

ALIGNMENTS

RESULT 1 AR043346

LOCUS AR043346 Sequence 23 from patent US 5814473.
DEFINITION Transamidases and aminotransferases
ACCESSION AR043346
VERSION AR043346.1 GI:5964354
KEYWORDS
SOURCE Unknown.

ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1065)
AUTHORS Warren,P.V. and Swanson,R.V.
TITLE Patent: US 5814473-A 23-29-SEP-1998;
FEATURES Location/Qualifiers

QY 601 GTAGTGTGAGACACTTCAAATGGATGGGAGTTAAGGTAGGGATTAA 660
Db 601 GTAGTGTGAGACACTTCAAATGGATGGGAGTTAAGGTAGGGATTAA 660
Db 661 GGGAGGGGAATCGCTCAGAACAAAGTCCTCACGAGCTCCCTCAAGTGACCTAC 720
Db 661 GGGAGGGGAATCGCTCAGAACAAAGTCCTCACGAGCTCCCTCAAGTGACCTAC 720
Qy 721 CCCTCAGTGATGCCAAAGTCCTCACGGAGGAAGAATGGAGAANG 780
Db 721 CCCTCAGTGATGCCAAAGTCCTCACGGAGGAAGAATGGAGAANG 780
Qy 781 ATACAGGAGTGTACAGCAGGAGAAGGATGACGACCAAGAAGAATGGAGAANG 780
Db 781 ATACAGGAGTGTACAGCAGGAGAAGGATGACGACCAAGAAGAATGGAGAANG 780
Qy 841 GTGAGGTTTCCGAGTAAGCTACTTGTGCTTGCAGAGCTAACCGGCCAC 900
Db 841 GTGAGGTTTCCGAGTAAGCTACTTGTGCTTGCAGAGCTAACCGGCCAC 900
Qy 901 GAGGTATAGGAGCTACTGAAAGGGATGCTCGTCAGGAACGATCTTACATGGAA 960
Db 901 GAGGTATAGGAGCTACTGAAAGGGATGCTCGTCAGGAACGATCTTACATGGAA 960
Qy 961 GGACTCAAAGTGCCTCAGGTAACTGAGGAAACCGGAGAACACAAGTTCTG 1020
Db 961 GGACTCAAAGTGCCTCAGGTAACTGAGGAAACCGGAGAACACAAGTTCTG 1020
Qy 1021 GAAGCACTGAGGAGATAAAACCCCTTCAGCTCTTTAA 1065
Db 1021 GAAGCACTGAGGAGATAAAACCCCTTCAGCTCTTTAA 1065

RESULT³
LOCUS AE000771/C
DEFINITION Aquifex aeolicus section 103 of 109 of the complete genome.
ACCESSION AE000771
VERSION AE000771.1
KEYWORDS .
SOURCE Aquifex aeolicus.
ORGANISM Aquifex aeolicus.
REFERENCE 1 (bases 1 to 12497)
AUTHORS Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E., Overbeek, R., Snead, M.A., Keller, M., AuJay, M., Huber, R., Feldman, R.A., Short, J.M., Olsen, G.J. and Swanson, R.V.
TITLE The complete genome of the hyperthermophilic bacterium Aquifex aeolicus
JOURNAL Nature 392 (6674), 353-358 (1998)
MEDLINE 98196666
PUBMED 9537320
REFERENCE 2 (bases 1 to 12497)
AUTHORS Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E., Overbeek, R., Snead, M.A., Keller, M., AuJay, M., Huber, R., Feldman, R.A., Short, J.M., Olsen, G.J. and Swanson, R.V.
TITLE JOURNAL
COMMENT Submitted (25-JUL-1997) Diversa Corporation, Genomics, San Diego, CA 92121
PUTATIVE indicates no similarity to known proteins
Hypothetical indicates similarity to a protein of unknown function.
Location/Qualifiers source
FEATURES
gene
gene_id="aq_2066"
organism="Aquifex aeolicus"
strain="VF5"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
gene
gene_id="aq_2073"
product="putative protein"
protein_id="AAC07818.1"

CDS
gene
gene_start=1
/codon_start=1
/transl_table=1
/product="hypothetical protein"
protein_id="AAC07821.1"
/db_xref="GI: 2984297"
/translation="MESIKETIEAFVPEGYGIGIFLFTAFESIIOQVPPDFITAGATA
FGNPLISALVATLNGTLLGGVAYFUGRFPVAKLGLGEKFKEAHTHRIGWA
VIAIGTPIPKVFCWLGSILELEVSRFTLASFIGRFPREFLFALEGHSIQGLF"
387-. 860
gene
gene_id="aq_2067"
/gene="aq_2067"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
protein_id="AAC07822.1"
/db_xref="GI: 2984298"
/translation="MIAFGNTTRSRSEGVYGFPLYROSEVFPLCPWVTLNSGFPKIFY
IMSFLVIVVSMTPLGKESVSYVHNLIDKSKYVLTPTMGTEEDWEMEVN
KGKFEALKEDCRISTIMKIDYRKGRGRLLQKVKSVQEKGRIODINATVEE"
934-. 1752
gene
gene_id="aq_2068"
note="aq_2068"
934-. 1752
gene
gene_id="argB"
/codon_start=1
/transl_table=1
/product="acetylglutamate kinase"
protein_id="AAC07812.1"
/db_xref="GI: 2984288"
/translation="MKYGSAMHDRELSFARDWVLUKKYVGINPVWVHGGPQISK
TEKFSTIRPKVGGMMRDEEMHVWEMVSGDINKNDIALNIRSGEKYIAYVLISGR
DGRLIAKKLDRKERYSELGIPVFDIGFGEVLTWNEELIFTISHNIPVAPVG
VGEEGRAYNVNADLASELAGEIKAKLTYLTDKGYLEBKLTISJSKDKAFLIK
KGVIROGMIPKVRSALRALERGVKKVHIDGRVKHSILLEVFTKEGVGTEITLE"
1765-. 2796
gene
gene_id="obg"
note="aq_2069"
1765-. 2796
CDS
gene
gene_id="obg"
/codon_start=1
/transl_table=1
/product="GTP-binding protein"
protein_id="AAC07816.1"
/db_xref="GI: 2984292"
/translation="MKFKVURVK_FWKGGGGDAVAFLREKYRKPGPAGGDGKGG
DVLIWTTSHLTLDEKKHHIAQNGEPORGKKNHGKOGEDLITYVPGTVVKDAQT
GEVICOLVKERQKCTVAKGGKGGRANARFPTPQNTAEEKGKGEWVILKELI
ADVGFGFPNAGSKSLTRAKKTAIYPETLSPNUGVMEDWLRVLLVIAKLP
IDEAHGAGLGEFLRHEIETKFLAHVIDVSDFRREPQVQAEFAINRELELPSKLAQ
KPQTIVVANKTDALSDRSLSLELEYKEKGYEFYAVSALTGEEGIEELKEGLWVKEET
RKDESQVT"
2822-. 3997
gene
gene_id="aq_2071"
CDS
gene
gene_id="aq_2071"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
protein_id="AAC07819.1"
/db_xref="GI: 2984295"
/translation="MLQVRKPGEDIRCFPPAWYRSEIASRSRKPKGETVVVROD
NSKFGGGYINPEVNAIRLTSDFKDNMPYMPPELRLRKTOAYVYKRYLINSNAYRUV
HSEGLDGLGIVLQVYKAVETFYGGVYVLFVQKPLVYKPLVYKPLVYKPLVYKPLVY
VEGLDVEWRYVGEPEENIWEHDKYVNIPIQGOKTGFFLDOARKFVRLVVEG
DRCLOVFCHGFFGFLAMKMRAGAGEVYDVSIELAKKEKAQNRLDCKWYKANAF
DYLRLDKKGEKPFVLLPSSPKFNRVPAVNAFLVPLVYKPLVYKPLVYKPLVY
CSFHTEQHLLDVLQASTDVRVORVRIAKTFQDLDHWIQLMENTLYLGWVVEV
3994-. 5151
gene
gene_id="aq_2073"
/codon_start=1
/transl_table=1
/product="putative protein"
protein_id="AAC07818.1"

/db-xref="GI: 2984294" gene
 /translation="MKHHWALVSGFNFQIPOSLIKVGGLTSLGDAHSLDA
 TASTLAFSLTSKFSERKEPFRYFGKLKLENIGATVAAFFLTAWNIIQALKGENIN
 FENLPIGIGVWLSLISLTSFLERBAGKLNSTLSDAHTDASSFLWILS
 SSIYFGNIEVRVAVAVALLVYTFELLEQIGAQLTSDAHTDASSFLWILS
 VSEVKRLLVNRAGKFDIDAVITNDIFKSHADETERKIMEKIPNDWVHYE
 PCCVKGTSVVALVAKOGVWALSFKVWDKIIIFKENEGNEPEVETGMDECIAKELAR
 RNDIVISGYHPTSNKAKWLHKNNVFWETEKENPYEELSEVSOLIKE
 5148 . 6395
 /gene="murd"
 /note="q_2075"
 5148 . 6395
 /gene="murd"
 /codon_start=1
 /transl_table=11
 /product="DOP-N-acetylumaramoylalanine-D-glutamate ligase"
 /protein_id="AAC07814_1"
 /ab_xref="GI: 2984290"
 /translation="MPLVWGGSSKGAKKLLKRGFEVAGDSQNPRLREVLGE
 VNVVUSPGTIPSPHPLWPKRKEKVEVGELEAVREFKGKVTAITGDKSTTRLT
 YLKKFDFEVAGNGKPSSEVLENPEGIAVLEVSPPGKTKTERNIGATSF
 SVDHLDWHPSTEDYLAKSKYRQNTEPDFTLNDLYEKTFRARKLSEYXIN
 SDSFYKQDILRNPKNUKIRGHNNVNASVHLTGLKPEDEEVIEFRGIPHR
 LEFLGNGVEVYNDNSSTPHALMHALKTPDNSVLTUGGKDGADEFSLRHVOK
 KVKIALAIGETERKIKSWDITEKINTLEAVKLAKEVRSKLSNWLFSPRCSSFD
 MFRVYEEGKTRKELVWFTRK"
 complement(6382 . 6996)
 /gene="trpF"
 /note="aq_2076"
 complement(6382 . 6996)
 /gene="trpF"
 /codon_start=1
 /transl_table=11
 /product="phosphoribosyl anthranilate isomerase"
 /protein_id="AAC07811_1"
 /ab_xref="GI: 2984287"
 /translation="MVKVKGITNTNDALECAKEAGADYIGVITYPKSPRTSKRKL
 ETKALEGIVKAVVAVNEPEFIKEILDEGFDLQHGDEDIEGKRVGDLKVKV
 FRYKEELTIVGKEMEKAVALDQFSKEAYGGKGTFFWEILKLVLEGHKVFLSGGLN
 PENKKEGTEFVKYGVDSVSGVEKEKKDFKVKREIERTKSE"
 7015 . 8556
 /gene="snf"
 /note="sq_2077"
 7015 . 8556
 /gene="snf"
 /codon_start=1
 /transl_table=11
 /product="(Na⁺):neurotransmitter symporter (Snf family)"
 /protein_id="AAC07817_1"
 /db_xref="GI: 2984293"
 /translation="MEVREHWATRGLLAMAGNAVGLGNFLRPPQOAENGGAFM
 IPTYIAFLVGLTIPMLWEMWAGVGGAGHGPATAYLWNRRFAITLGPGWLPL
 WATIYVWESWFLGFAIKELVGLVPPNMPDPSLSPKPELIVSYIYPKGP
 ILKRSLFATVFLTMIINSTVLTGKIERFKAMPFLFLAVLFLVIRFLV
 PNSTGADGUNFLWDPFLKIRGKDGWTAVGFLVIRFLV
 LSGITAATUNKEEVILGGSISIPAAVFGVANAVATAKAGAFNIGFITAESOT
 AGGFGLGFWMFLFLFFAGTTSIIMOPRFLDEKLRSKRHVAWLTAVFSAHL
 VMFLNKSLSDEMDFDPAFTGCVVFLGFLTTEIFVGDAKAWBENRGLKIPRYY
 VMRLITPAFLAVLWVWAREYLPKIMEETHWTVITRFLYIGLFLFLTVFLAERR
 NHESR
 8543 . 9394
 /gene="sppA"
 /note="sq_2080"
 8543 . 9394
 /gene="sppA"
 /codon_start=1
 /transl_table=11
 /product="protease IV"
 /protein_id="AAC07815_1"
 /db_xref="GI: 2984291"
 /translation="MRVWKWLITFLSLLIFLSVKEFILSKFPVGHJLAVLKGSVTV
 NPFLVNLKFLERKUSKQVAVLWRVSDGSGVSAQOEYRAFERSKGPKVHSN
 VAAGGGYVYSAAGYKIFANEGETITGSGIVIOWAYKLVKEGLIETAIKGKFDT

/note="Pfam match to entry aminotran_1, Aminotransferase class-I, score -25.9, E-value 1.60E-08, complement(3478..3482); /gene="HisB"

CDS

```

complement(2255..2866)
/gene="HisH"
/note="Best Blastp hit = gi|13913863|sp|066943|HIS5_AQUAE
AMIDOTRANSFERASE_HISH gi|74339861|pir|IC70364
amidotransferase_HISH - Aquifex aeolicus
91|2983311|pir|AAC06901.1|AE000705) amidotransferase_HISH
[Aquifex aeolicus], score 207, E-value 7.00E-53"
/codon_start=1
/translate_table=11
/protein_id="AAM25300.1"
/ab_xref="GI:20517148"
/translation=MIGIDIGGNLRSVEKALKIYVGDAKTSDAKEIKWELLIP
GVSAFPDAIRVLLKTMGAEIKRHKIEKPKFGLQHQLIFEKSFPGKTYGLKIFK
GIVVGLPKNKTRPIKGKTPPLKGKTFVIVSYVYVNGNEKFEEAVT
EGIEPATIKDNVFSQFHPKSGETGLKLKNFGEMIKL"
complement(2267..2860)
```

misc_feature

```

/gene="HisH"
/note="Pfam match to entry GATase, Glutamine amidotransferase class-I, score 176.7, E-value 3.80E-49"
complement(2275..2279)
/gene="HisA"
complement(2874..2878)
```

gene

```

/gene="HisB"
complement(2885..3482)
```

RBS

```

/RBS
/gene="HisB"
complement(2821..369)
```

CDS

```

/codon_start=1
/translate_table=11
/product="Imidazoleglycerol-phosphate dehydratase"
/protein_id="AM25301.1"
/ab_xref="GI:20517149"
/translation=MRAEKVKRKTAEDTYVLENDIGKGSYDITKGIGFFDHMLSLFA
g1|2398791|emb|CA60712.1| (X87266) imidazoleglycerol-phosphate dehydratase [Rhodobacter sphaeroides], score 223, E-value 2.00E-57"
/codon_start=1
/translate_table=11
/product="Imidazoleglycerol-phosphate dehydratase"
/protein_id="AM25301.1"
/ab_xref="GI:20517149"
/translation=MRAEKVKRKTAEDTYVLENDIGKGSYDITKGIGFFDHMLSLFA
KHGDIDLVIAKGIDEVHDHTVLDVGLTVALAKASGDKKKIKREFTFYVMDAL
VRVSUDISGRPLYIDLPKAERVQNFETEVFRFAYNQGITHVELLHGYNTH
HIRESPKAIGKALDEALKIDERIDGIPSTKIL"
complement(2245..3379)
```

misc_feature

```

/gene="HisB"
/note="Pfam match to entry IGPD, imidazoleglycerol-phosphate dehydratase, score 293.4, E-value 6.20E-100"
complement(3466..4534)
```

gene

```

/gene="HisC2"
/note="Pfam match to entry IGPD, imidazoleglycerol-phosphate dehydratase, score 293.4, E-value 6.20E-100"
complement(3466..4521)
```

CDS

```

/gene="HisC2"
/note="Best Blastp hit = gi|17450424|pir|IET0478 histidinol-phosphate amidotransferase - Aquifex aeolicus gi|28842891|gb|AAC07813.1| (AE000711) histidinol-phosphate amidotransferase [Aquifex aeolicus], score 230, E-value 2.00E-59"
/codon_start=1
/translate_table=11
/product="Histidinol-phosphate amidotransferase/tyrosine aminotransferase"
/protein_id="AM25302.1"
/ab_xref="GI:20517150"
/translation=MENLILREKEIKGFKNEYENVNPYKMDANETPFELPEVMKNI
GDIKSHVHNYPDPKAERELARYCSYTPKNTFVGMGSDTEIHLMLAFVVKGD
VLYPHSFAMISITSIKAGANEIYVANLNEFTINTERFAVERKPKVPCPNP
TGSVIDBEDILITIERTKARTVIVTDEAYFEVYKTUVPYDREFENITVLRISKA
GIRGVGALSNETVKVLNLYKSPYNNSLSORIALEVLGSKVLRVNYLNEREKL
KBLNKINGLKPSPHANEFLCKFENANDVHVRNLVERGLVLRNFSWKLEGTLRIVS
SSDANKLNALREILS"
complement(3469..4515)
```

misc_feature

```

/gene="HisC2"
/complement(3469..4515)
```


DKIHTVLDGAFSGKPNOLIDIVEKLTIRDGVVRVQGGFRDVYESIKAVDGINENVII
GTKALDSFLERLINSFEGFTIVSLSKGKGRFTKGWIERSLVVEDAYMLRKVNRF
VTVIDRGDTIGIEKIKRWGDEEIFIYAGGAATPTDIOLAEGFSGVIGKALVEG
TISLERALEVAKSVGKKNYSSS" 4570. . 5370

gene "PFI663"
CDS
/gene="PFI663"
/note="Function Code: 1.7 Amino Acid Biosynthesis:
Histidine family; (HIS_F)"
/codon_start=1
/transl_table=11
/product="imidazoleglycerol-phosphate synthase, cyclase"
subunit"
/protein_id="AAL81787_1"
/db_xref="GI:18893820"
/translation="MKEELSLKLARLLKVLRARRITAIADIKDSGRVKGSIKFKNRDA
GDPIELAKRVESEGIDEIVTLTASYEKILUVERAETIYVFVPGGJRTVE
ERIRRGADKIFINTAAENPSLKEADLGSANLYVAIDAKWNGKFWEWYTHGG
RKPRGDAVEKAKRVEELGAGEELLTSMDGDKTGFDPLIKAVDANIPIVASSG
5345.. 6010
gene "PFI664"
CDS
/gene="PFI664"
/note="Function Code: 1.7 Amino Acid Biosynthesis:
Histidine family"
/codon_start=1
/transl_table=11
/product="phosphoribosyl-AMP cyclohydrolase /
phosphoribosyl-atp"
/protein_id="AAL81788_1"
/db_xref="GI:18893821"
/translation="MFQDUDRGEHMEIPLLQEONWERNNGTIVPVTDKEGVLT
LAVMNEALRTLEPLGIAHYHSRSKTRKMGEVNSNTQVKWIKDCNDALLIVS
PKGPACTIGNISCFFRKLGEBERVLIDSYSLKELLEIKRKKETPVGSTSKLF
RKGREKLYKKFGEAEATLEVLAKEKRERLIVYEVDLLYLVLTYNDISLGEMNLRR
RRK"
6007. . 7023
gene "PFI665"
CDS
/gene="PFI665"
/note="Function Code: 1.7 Amino Acid Biosynthesis:
Histidine family"
/codon_start=1
/transl_table=11
/product="putative histidinol-phosphate aminotransferase
(imidazole"
/protein_id="AAL81789_1"
/db_xref="GI:18893822"
/translation="MWEETLNFEPYRAVEGNRYIWLDKKNESPYDLPOLKEELEL
KRIEFNPKPHTSDFPDLREALFEYGKJENATAVGNGSDELINILVKMKRITIVTSP
FGCMYSPFAKLGIPYKDPKIPKEFDIDGERIAEKGAKASAFIASPNNTGNSOPEN
EVLKVLDSGRWVILDEAYSESGSNSPKTPIKSEYENENVLUTFSKAFLAGTICGMIA
NERKIDBLRKLPPINSLNTMVAKMELNDIYKRLILVKERERREPEFYSY
PSEANFLIMKLQDAYDYLKKGIVVRKLSLRGLRHGTICKWWENDELIKALKEFLEE
CRCCG"
gene "PFI666"
CDS
/gene="PFI666"
/note="Function Code: 16.1 Conserved Hypothetical"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 9.7%; Score 143.8; DB: 1; Length 12452;
Best Local Similarity 51.9%; Pred. No. 4.7e-18;
Matches 200; Conservative 0; Mismatches 257; Indels 12; Gaps 2;
QY 65 GGCTTCTCAAGAACATTCCCTACGACTTCCGAGGAGATAAACAAAGGCCCTAG 124
Db 6071 GGCTTCTCAAGAACATTCCCTACGACTTCCGAGGAGATAAACAAAGGCCCTAG 6130

RESULT 6

REFERENCE	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Neisseria meningitidis	AE002508	1			Neisseria meningitidis MC58
Neisseria, Proteobacteria, beta subdivision: Neisseriaceae; Neisseria.					
1 (basses 1 to 11932)					
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Cittone, H., Clark, E.B., Cottont, M.D., Utterback, T.R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scariato, V., Masiognani, V., Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappoport, R. and Venter, J.C.					
Complete genome sequence of Neisseria meningitidis serogroup B strain MC58					
Science 287 (5459), 1809-1815 (2000)					
JOURNAL	10710307	2017575	(bases 1 to 11932)	PUBMED	
ARTICLE	Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Cittone, H., Clark, E.B., Cottont, M.D., Utterback, T.R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scariato, V., Masiognani, V.,			AUTHORS	REFERENCE

RAPPOLI, R. and VENTER, J.C. *Rippa, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappoli, R. and Venter, J.C.*
 TITLE
 JOURNAL
 Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 FEATURES
 source
 gene
 CDS
 /organism="Neisseria meningitidis MC58"
 /strain="MC58"
 /db_xref="taxon:122886"
 /note="Serogroup: B"
 102..1097
 /gene="NMB1573"
 102..1097
 /gene="NMB1573"
 /note="similar to GB:M34930 SP:P21302 GB:X64860 PID:150241
 PID:44847 Percent identity: 97.58; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="ornithine carbamoyltransferase, catabolic"
 /protein_id="AAF41926_1"
 /db_xref="GI:7226821"
 /translation="MNLKRNHFLKLIDFTPEETIAYDLAELKAKKACREIORMKG
 KNALIFEFKSTRTRCAFEVARDGAGVYVLEPSASOIQGHKESTDTARRLGRMDA
 IERFGCQEVEELAKYAGPVGNGTNEHPTOMLADAJMRENSKPINOTAEVY
 GDARKNGTNSILLIGLAKGDVRGAPOSWPSCTIAAHHAAKETGAITLTENAH
 EAVKNDFTDWDWSMGEKEYWQERIDLDKDYKTPPEMASCNPQKFMCILPAP
 HNRKTKGENIYEFGLNGEVTEELFESPASIVFDQAERNHHTKAVMVLAQD"
 /gene="NMB1574"
 complement("1321..2334")
 /gene="NMB1574"
 /note="similar to GB:AE000512 percent identity: 77.25;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="ketol-acid reductoisomerase"
 /protein_id="AAF41927_1"
 /db_xref="GI:7226822"
 /translation="MQVYTKDADLSLIKKTVAATIGYGSQGHAAHNLLDSGVNVT
 GLRGCGSSWKKAEGAGHVVYKVAEAKDQVWMLLPDEPMPPAVHAEVYANLKGATL
 AFAHGDNVHNQTVPRDIDVIMARPGGHTVSEYIKRGGGVPSSLIAVYQDNGSKAK
 DIAUSYAAGGGTTRGVIENTFREETDGFEGQAVLCCGGVVERLKGAGTETRAGYA
 PEMAFECHEMKLKVLDLFFEGGTTANMTSISNNHVEYGEVUTGGEVUNMASKEAMRA
 LKRQTGEYAMFOQEGNNTVASYMSTARRLNLADHOVERVKAQQLRAMPWTANKLVQ
 DN"
 complement("2414..2707")
 /gene="NMB1575"
 complement("2414..2707")
 /note="conserved hypothetical protein; identified by
 Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="NMB1575"
 /db_xref="GI:7226823"
 /translation="MSNTRKVALVTPVKPYETLAAOFRELVKRASRAEEGNTSYDLHQ
 EIGKPNRFEVEMNKSOAIDEHNAHRCQAFVOSVDGKTEALEITVLMNEAVD"
 complement("2774..3265")
 /gene="NMB1576"
 /note="similar to DBR:147816 GB:U00096 PID:1786266 percent
 identity: 79.75; identified by sequence similarity;
 putative"
 /codon_start=1
 /transl_table=11
 /product="acetolactate synthase III, small subunit"
 /protein_id="AAF41929_1"
 /db_xref="GI:7226824"
 /note="similar to DBR:147816 GB:U00096 PID:1786266 percent
 identity: 79.75; identified by sequence similarity;
 putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAF41933_1"
 /db_xref="GI:7226828"
 /translation="MRHLISLVIENESSAMSRVGLSARDNIDSLSLAVPTEDKTL"

gene	NRIFEHENWVASRQDPVSGETIDTRKLKATSKCKRLGRYITSYHLEVEALAAEALTHFVKSYLGKKLKNKLTAGHOPPGISVLEEDRRFGSSLPLNIEPDGLIAKHNARDLDTDEELDYLFDNDODDANKSRTGNGNWIGKDVIAKVKVIVGHGLONAVSAKEDGFETAVETETNGRQSRYRPTTSEEVILAKLGLQQKCLPELKFGSGKEAYIRPKTEPEQENTOTTPNPKIKRETKS"	RESULT	7	AX044033	Sequence	349980 bp	DNA	linear	PAT	24 -NOV -2000
CDS	8509 .. 9798	ACCESSION	AX044033	VERSION	AX044033.1	GI	11342917	SOURCE	Neisseria meningitidis.	
/gene="NMB1581"	/note="similar to GB:AB000657 percent identity: 71.03; identified by sequence similarity; putative"	/codon_start=1	/transl_table=11	/product="histidinol dehydrogenase"	/protein_id="AAK41934.1"	/db_xref="GI:226839"	/translation="MKKINTQSDFOASIKALLAFETAQNPERTIVADICAVDQKRGDALEYTKFEDQDNKSTDDLNALNARVYQHORQKMEHSVYTDQDGTLQQTPLDRVGVPGGKRAVPSVIMMAPVAVGKREIMVPTPGERNDIVLAAYAVGVTYFVSGAQAVALAQTETPPQVKITGCAFAVAAKKRKYFGVGILDIVAGSEILVFTADGTPADWVAMLDLSQAHDELQAQALIGTSQA			
Query Match 8.5%; Score 90.8; DB 1; Length 11932; Best Local Similarity 47.7%; Pred. No. 2.e-14; Matches 331; Conservative 0; Mismatches 357; Indels 6; Gaps 2; Matches 331; Conservative 0; Mismatches 357; Indels 6; Gaps 2;	Qy 182 AAGGGTCTCGGGAATTTCGCGCTTAGGAAGAAATTAGTTAGTCGGTACGGT 241	Db 10058 AACATTAACGTTCGCGTCCGACTATCCCGGACTGGGCCGACATCGCGCTGGCACGGT 10117								
Qy 242 CGGACGACTCATATCCTCTCATAGCTTACACCCCTTACA 301	Db 10118 CGGACGACTGATACGTTCACACCGTGCAGCACCGCCAAACCGGGCGAACATGTTG 10177									
Qy 302 TACCTGTCCGACCTTCCATGAGATAAGCGGAAGTCTCGGAGAACCTCG 361	Db 10178 CAGCCGACACCGTTCAGTCATGCCACAGCCGCGCTGATGGATTAG 10237									
Qy 362 TAAAGGTCAACTGGAGAACCTTGTATAGACTTAGAAGAAGTATGAATTAATAG 421	Db 10238 TCGCGCTTCCACTGACCGGAGATTCACCCCAACTGCGCCCTCTCGAACCGCTCA 10297									
Qy 422 AGAAAGAAAACCCTGCTCGGGTACTTTGTTACCCCAACACGGAAACCTCT 481	Db 10298 GGAACACCGCCCTGCGCTGACCTTATCGCTTACCCCAACACGGCGTATGCT 10357									
Qy 482 TTTCAGGGGAAGATGA--GGGATAGAACAGGGGTTCTCTGTAATAGCG 538	Db 10358 TCAAGCTGCGCAATCGAAGCGCTCATCGTACAGGCACTGCTGATG 10417									
Qy 539 AACGCTACTATCACTCCGGAAACCTTCTGGAGAGCGCTCAAAGGGAGATA 598	Db 10418 AACCTACTGGCCATTCAACGGCGACAGCTTCTGGCAGGGCAGATTCAC 10477									
Qy 599 CGCTAGTTTGAGACACTTCAAAATCGGTTAGGGTAGGGATTAA 658	Db 10478 TGATAGTCTTACGCCCTCAGCAAAATCGTTTGGCGGACTGCCATCGGTATGGCG 10537									
Qy 659 TAGGAAAGGGGAATGTCCTCAGAAATTACAGGTGAGACTCCCTCTACGGTACCT 718	Db 10538 CAGCTGCCCGAAGTCATCGGGACTGGGTTAACCGCCATATGAAACC 10597									
Qy 719 ACCCTCTCGGTGATGCCAAAGTCTCTCAGGGAGAGATTCTCTATGGAA 778	Db 10598 AATGAGCTGACACTGCCAACCTCCCTGGGACTAGGCCATTCTCT--GCCA 10654									
Qy 779 AGTACAGGAGGTGTCAGACAGGAGAAAGTGTGAGCAATGAGAAATGAG 838	Db 10655 ACATGACAGCCTGAAACACGAGCGACGGTGTGCGGCAATTGGCAAAATGCC 10714									
Qy 839 GAGTGTGAGTTTCGCTGAGTAAAGCTCTCT 872	Db 10715 GTCTGAAACACCTTTCAAGTCAGGCAACTCAT 10748									
REFERENCE 1 (bases 1 to 349980) AUTHORS Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J.C., Masignani, V., Galleotti, C., Mora, M., Ratti, G., Sarselli, M., Scariato, V., Rappuoli, R., Fraser, C.M. and Grandi, G. TITLE Neisseria genomic sequences and methods of their use PATENT: WO 0066791-A-112 09-NOV-2000; CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US) JOURNAL Location/Qualifiers										
FEATURES source /organism="Neisseria meningitidis" /ab_xref="taxon:487" /note="sequence too long, cut in 8 pieces..seq 1 to 349980 bases..seq 108: 300001 to 649980 349980 bases..seq 110: 900001 to 1249980 349980 bases..seq 111: 1200001 to 1549980 349980 bases..seq 112: 180001 to 2149980 349980 bases..seq 114: 210001 to 2272325 172325 bases..seq 115: 2272325 172325 bases..seq 116: 2272325 172325 bases..seq 117: 2272325 172325 bases..seq 118: 2272325 172325 bases..seq 119: 2272325 172325 bases..seq 120: 2272325 172325 bases..seq 121: 2272325 172325 bases..seq 122: 2272325 172325 bases..seq 123: 2272325 172325 bases..seq 124: 2272325 172325 bases..seq 125: 2272325 172325 bases..seq 126: 2272325 172325 bases..seq 127: 2272325 172325 bases..seq 128: 2272325 172325 bases..seq 129: 2272325 172325 bases..seq 130: 2272325 172325 bases..seq 131: 2272325 172325 bases..seq 132: 2272325 172325 bases..seq 133: 2272325 172325 bases..seq 134: 2272325 172325 bases..seq 135: 2272325 172325 bases..seq 136: 2272325 172325 bases..seq 137: 2272325 172325 bases..seq 138: 2272325 172325 bases..seq 139: 2272325 172325 bases..seq 140: 2272325 172325 bases..seq 141: 2272325 172325 bases..seq 142: 2272325 172325 bases..seq 143: 2272325 172325 bases..seq 144: 2272325 172325 bases..seq 145: 2272325 172325 bases..seq 146: 2272325 172325 bases..seq 147: 2272325 172325 bases..seq 148: 2272325 172325 bases..seq 149: 2272325 172325 bases..seq 150: 2272325 172325 bases..seq 151: 2272325 172325 bases..seq 152: 2272325 172325 bases..seq 153: 2272325 172325 bases..seq 154: 2272325 172325 bases..seq 155: 2272325 172325 bases..seq 156: 2272325 172325 bases..seq 157: 2272325 172325 bases..seq 158: 2272325 172325 bases..seq 159: 2272325 172325 bases..seq 160: 2272325 172325 bases..seq 161: 2272325 172325 bases..seq 162: 2272325 172325 bases..seq 163: 2272325 172325 bases..seq 164: 2272325 172325 bases..seq 165: 2272325 172325 bases..seq 166: 2272325 172325 bases..seq 167: 2272325 172325 bases..seq 168: 2272325 172325 bases..seq 169: 2272325 172325 bases..seq 170: 2272325 172325 bases..seq 171: 2272325 172325 bases..seq 172: 2272325 172325 bases..seq 173: 2272325 172325 bases..seq 174: 2272325 172325 bases..seq 175: 2272325 172325 bases..seq 176: 2272325 172325 bases..seq 177: 2272325 172325 bases..seq 178: 2272325 172325 bases..seq 179: 2272325 172325 bases..seq 180: 2272325 172325 bases..seq 181: 2272325 172325 bases..seq 182: 2272325 172325 bases..seq 183: 2272325 172325 bases..seq 184: 2272325 172325 bases..seq 185: 2272325 172325 bases..seq 186: 2272325 172325 bases..seq 187: 2272325 172325 bases..seq 188: 2272325 172325 bases..seq 189: 2272325 172325 bases..seq 190: 2272325 172325 bases..seq 191: 2272325 172325 bases..seq 192: 2272325 172325 bases..seq 193: 2272325 172325 bases..seq 194: 2272325 172325 bases..seq 195: 2272325 172325 bases..seq 196: 2272325 172325 bases..seq 197: 2272325 172325 bases..seq 198: 2272325 172325 bases..seq 199: 2272325 172325 bases..seq 200: 2272325 172325 bases..seq 201: 2272325 172325 bases..seq 202: 2272325 172325 bases..seq 203: 2272325 172325 bases..seq 204: 2272325 172325 bases..seq 205: 2272325 172325 bases..seq 206: 2272325 172325 bases..seq 207: 2272325 172325 bases..seq 208: 2272325 172325 bases..seq 209: 2272325 172325 bases..seq 210: 2272325 172325 bases..seq 211: 2272325 172325 bases..seq 212: 2272325 172325 bases..seq 213: 2272325 172325 bases..seq 214: 2272325 172325 bases..seq 215: 2272325 172325 bases..seq 216: 2272325 172325 bases..seq 217: 2272325 172325 bases..seq 218: 2272325 172325 bases..seq 219: 2272325 172325 bases..seq 220: 2272325 172325 bases..seq 221: 2272325 172325 bases..seq 222: 2272325 172325 bases..seq 223: 2272325 172325 bases..seq 224: 2272325 172325 bases..seq 225: 2272325 172325 bases..seq 226: 2272325 172325 bases..seq 227: 2272325 172325 bases..seq 228: 2272325 172325 bases..seq 229: 2272325 172325 bases..seq 230: 2272325 172325 bases..seq 231: 2272325 172325 bases..seq 232: 2272325 172325 bases..seq 233: 2272325 172325 bases..seq 234: 2272325 172325 bases..seq 235: 2272325 172325 bases..seq 236: 2272325 172325 bases..seq 237: 2272325 172325 bases..seq 238: 2272325 172325 bases..seq 239: 2272325 172325 bases..seq 240: 2272325 172325 bases..seq 241: 2272325 172325 bases..seq 242: 2272325 172325 bases..seq 243: 2272325 172325 bases..seq 244: 2272325 172325 bases..seq 245: 2272325 172325 bases..seq 246: 2272325 172325 bases..seq 247: 2272325 172325 bases..seq 248: 2272325 172325 bases..seq 249: 2272325 172325 bases..seq 250: 2272325 172325 bases..seq 251: 2272325 172325 bases..seq 252: 2272325 172325 bases..seq 253: 2272325 172325 bases..seq 254: 2272325 172325 bases..seq 255: 2272325 172325 bases..seq 256: 2272325 172325 bases..seq 257: 2272325 172325 bases..seq 258: 2272325 172325 bases..seq 259: 2272325 172325 bases..seq 260: 2272325 172325 bases..seq 261: 2272325 172325 bases..seq 262: 2272325 172325 bases..seq 263: 2272325 172325 bases..seq 264: 2272325 172325 bases..seq 265: 2272325 172325 bases..seq 266: 2272325 172325 bases..seq 267: 2272325 172325 bases..seq 268: 2272325 172325 bases..seq 269: 2272325 172325 bases..seq 270: 2272325 172325 bases..seq 271: 2272325 172325 bases..seq 272: 2272325 172325 bases..seq 273: 2272325 172325 bases..seq 274: 2272325 172325 bases..seq 275: 2272325 172325 bases..seq 276: 2272325 172325 bases..seq 277: 2272325 172325 bases..seq 278: 2272325 172325 bases..seq 279: 2272325 172325 bases..seq 280: 2272325 172325 bases..seq 281: 2272325 172325 bases..seq 282: 2272325 172325 bases..seq 283: 2272325 172325 bases..seq 284: 2272325 172325 bases..seq 285: 2272325 172325 bases..seq 286: 2272325 172325 bases..seq 287: 2272325 172325 bases..seq 288: 2272325 172325 bases..seq 289: 2272325 172325 bases..seq 290: 2272325 172325 bases..seq 291: 2272325 172325 bases..seq 292: 2272325 172325 bases..seq 293: 2272325 172325 bases..seq 294: 2272325 172325 bases..seq 295: 2272325 172325 bases..seq 296: 2272325 172325 bases..seq 297: 2272325 172325 bases..seq 298: 2272325 172325 bases..seq 299: 2272325 172325 bases..seq 300: 2272325 172325 bases..seq 301: 2272325 172325 bases..seq 302: 2272325 172325 bases..seq 303: 2272325 172325 bases..seq 304: 2272325 172325 bases..seq 305: 2272325 172325 bases..seq 306: 2272325 172325 bases..seq 307: 2272325 172325 bases..seq 308: 2272325 172325 bases..seq 309: 2272325 172325 bases..seq 310: 2272325 172325 bases..seq 311: 2272325 172325 bases..seq 312: 2272325 172325 bases..seq 313: 2272325 172325 bases..seq 314: 2272325 172325 bases..seq 315: 2272325 172325 bases..seq 316: 2272325 172325 bases..seq 317: 2272325 172325 bases..seq 318: 2272325 172325 bases..seq 319: 2272325 172325 bases..seq 320: 2272325 172325 bases..seq 321: 2272325 172325 bases..seq 322: 2272325 172325 bases..seq 323: 2272325 172325 bases..seq 324: 2272325 172325 bases..seq 325: 2272325 172325 bases..seq 326: 2272325 172325 bases..seq 327: 2272325 172325 bases..seq 328: 2272325 172325 bases..seq 329: 2272325 172325 bases..seq 330: 2272325 172325 bases..seq 331: 2272325 172325 bases..seq 332: 2272325 172325 bases..seq 333: 2272325 172325 bases..seq 334: 2272325 172325 bases..seq 335: 2272325 172325 bases..seq 336: 2272325 172325 bases..seq 337: 2272325 172325 bases..seq 338: 2272325 172325 bases..seq 339: 2272325 172325 bases..seq 340: 2272325 172325 bases..seq 341: 2272325 172325 bases..seq 342: 2272325 172325 bases..seq 343: 2272325 172325 bases..seq 344: 2272325 172325 bases..seq 345: 2272325 172325 bases..seq 346: 2272325 172325 bases..seq 347: 2272325 172325 bases..seq 348: 2272325 172325 bases..seq 349: 2272325 172325 bases..seq 350: 2272325 172325 bases..seq 351: 2272325 172325 bases..seq 352: 2272325 172325 bases..seq 353: 2272325 172325 bases..seq 354: 2272325 172325 bases..seq 355: 2272325 172325 bases..seq 356: 2272325 172325 bases..seq 357: 2272325 172325 bases..seq 358: 2272325 172325 bases..seq 359: 2272325 172325 bases..seq 360: 2272325 172325 bases..seq 361: 2272325 172325 bases..seq 362: 2272325 172325 bases..seq 363: 2272325 172325 bases..seq 364: 2272325 172325 bases..seq 365: 2272325 172325 bases..seq 366: 2272325 172325 bases..seq 367: 2272325 172325 bases..seq 368: 2272325 172325 bases..seq 369: 2272325 172325 bases..seq 370: 2272325 172325 bases..seq 371: 2272325 172325 bases..seq 372: 2272325 172325 bases..seq 373: 2272325 172325 bases..seq 374: 2272325 172325 bases..seq 375: 2272325 172325 bases..seq 376: 2272325 172325 bases..seq 377: 2272325 172325 bases..seq 378: 2272325 172325 bases..seq 379: 2272325 172325 bases..seq 380: 2272325 172325 bases..seq 381: 2272325 172325 bases..seq 382: 2272325 172325 bases..seq 383: 2272325 172325 bases..seq 384: 2272325 172325 bases..seq 385: 2272325 172325 bases..seq 386: 2272325 172325 bases..seq 387: 2272325 172325 bases..seq 388: 2272325 172325 bases..seq 389: 2272325 172325 bases..seq 390: 2272325 172325 bases..seq 391: 2272325 172325 bases..seq 392: 2272325 172325 bases..seq 393: 2272325 172325 bases..seq 394: 2272325 172325 bases..seq 395: 2272325 172325 bases..seq 396: 2272325 172325 bases..seq 397: 2272325 172325 bases..seq 398: 2272325 172325 bases..seq 399: 2272325 172325 bases..seq 400: 2272325 172325 bases..seq 401: 2272325 172325 bases..seq 402: 2272325 172325 bases..seq 403: 2272325 172325 bases..seq 404: 2272325 172325 bases..seq 405: 2272325 172325 bases..seq 406: 2272325 172325 bases..seq 407: 2272325 172325 bases..seq 408: 2272325 172325 bases..seq 409: 2272325 172325 bases..seq 410: 2272325 172325 bases..seq 411: 2272325 172325 bases..seq 412: 2272325 172325 bases..seq 413: 2272325 172325 bases..seq 414: 2272325 172325 bases..seq 415: 2272325 172325 bases..seq 416: 2272325 172325 bases..seq 417: 2272325 172325 bases..seq 418: 2272325 172325 bases..seq 419: 2272325 172325 bases..seq 420: 2272325 172325 bases..seq 421: 2272325 172325 bases..seq 422: 2272325 172325 bases..seq 423: 2272325 172325 bases..seq 424: 2272325 172325 bases..seq 425: 2272325 172325 bases..seq 426: 2272325 172325 bases..seq 427: 2272325 172325 bases..seq 428: 2272325 172325 bases..seq 429: 2272325 172325 bases..seq 430: 2272325 172325 bases..seq 431: 2272325 172325 bases..seq 432: 2272325 172325 bases..seq 433: 2272325 172325 bases..seq 434: 2272325 172325 bases..seq 435: 2272325 172325 bases..seq 436: 2272325 172325 bases..seq 437: 2272325 172325 bases..seq 438: 2272325 172325 bases..seq 439: 2272325 172325 bases..seq 440: 2272325 172325 bases..seq 441: 2272325 172325 bases..seq 442: 2272325 172325 bases..seq 443: 2272325 172325 bases..seq 444: 2272325 172325 bases..seq 445: 2272325 172325 bases..seq 446: 2272325 172325 bases..seq 447: 2272325 172325 bases..seq 448: 2272325 172325 bases..seq 449: 2272325 172325 bases..seq 450: 2272325 172325 bases..seq 451: 2272325 172325 bases..seq 452: 2272325 172325 bases..seq 453: 2272325 172325 bases..seq 454: 2272325 172325 bases..seq 455: 2272325 172325 bases..seq 456: 2272325 172325 bases..seq 457: 2272325 172325 bases..seq 458: 2272325 172325 bases..seq 459: 2272325 172325 bases..seq 460: 2272325 172325 bases..seq 461: 2272325 172325 bases..seq 462: 2272325 172325 bases..seq 463: 2272325 172325 bases..seq 464: 2272325 172325 bases..seq 465: 2272325 172325 bases..seq 466: 2272325 172325 bases..seq 467: 2272325 172325 bases..seq 468: 2272325 172325 bases..seq 469: 2272325 172325 bases..seq 470: 2272325 172325 bases..seq 471: 2272325 172325 bases..seq 472: 2272325 172325 bases..seq 473: 2272325 172325 bases..seq 474: 2272325 172325 bases..seq 475: 2272325 172325 bases..seq 476: 2272325 172325 bases..seq 477: 2272325 172325 bases..seq 478: 2272325 172325 bases..seq 479: 2272325 172325 bases..seq 480: 2272325 172325 bases..seq 481: 2272325 172325 bases..seq 482: 2272325 172325 bases..seq 483: 2272325										

QY 719 ACCCTCTCAGGTGATGGCAAAAGTCTCTCACGGAGGAAGAATTCCATGGAA 778
 | || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 145159 AATGAGGCTGACACTCCCAACTCGCCCTCGGCACATTCCTCTT---GCCA 145215
 QY 779 AGATACAGGAGGTGTAACAGAGGAGAAAGGATGTTACGACCAAAGAAGAAAATGAG 838
 | || | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 145216 ACATCGACGCTGAAAAACGACGCGAACGGATGTTGCCAATTGGCAATAATGCC 145275
 QY 839 GAGTGAGGTTTCGAGTAGGCTACTCTT 872
 | || | || | || | || | || | || | || | || | || | || | || | || |
 Db 145276 GCTGAGCACCTTTCAGTCAGTGGCAACTCTAT 145309

RESULT 8
 A001764/c

LOCUS AE001764 20334 bp DNA linear BCT 02-JUN-1999
DEFINITION Thermotoga maritima section 76 of 136 of the complete genome.
ACCESSION AE001764 AE000512
VERSION AE001764.1 GI:4981561
KEYWORDS: thermotoga maritima
SOURCE: Thermotoga maritima
ORGANISM: Thermotoga maritima
BACTERIA: Thermotogae; Thermotogales; Thermotogaceae;
PHYLUM: Thermotoga
CLASS: I to 20334
REFERENCE:
 1. (bases 1 to 20334)
AUTHORS: Nelson, K.E., Clayton, R.A., Gill, S.R., Gwynn, M.L., Dodson, R.J.,
 Haft, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K.A.,
 McDonald, L., Utterback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M.,
 Stewart, A.M., Cotton, M.D., Pratt, M.S., Phillips, C.A., Richardson, D., Heidelberg, J., Sutton, G., Fleischmann, R.D., Eisen, J.A., Fraser, C.M. et al.

TITLE: Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima
JOURNAL: Nature 399 (6734), 323-329 (1999)
MEDLINE: 99287316
PUBMED: 10360571

REFERENCE:
 2. (bases 1 to 20334)
AUTHORS: Nelson, K.E., Clayton, R.A., Gill, S.R., Gwynn, M.L., Dodson, R.J.,
 Haft, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K.A.,
 McDonald, L., Utterback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M.,
 Stewart, A.M., Cotton, M.D., Pratt, M.S., Phillips, C.A., Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D., White, O., Salzberg, S.L., Smith, H.O., Venter, J.C. and Fraser, C.M.

TITLE: Direct Submission
JOURNAL: Submitted (01-JUN-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES:
 source
 1. . 20334
 /organism="Thermotoga maritima"
 gene
 /db_xref="taxon:2336"
 complement("70", 1338)
 /gene="TM1021"
 complement("70", 1338)
 /gene="TM1021"
 /note="similar to GP:2695718 percent identity: 100.00;
 /codon_start=1
 /transl_table=11
 /product="permease, putative"
 /protein_id="AA336098_1"
 /db_xref="GI:4981562"
 /translation="MAVITILLMLVILNADOMWMSPNIGALEQEFNITDAQIQLVASS
 FTVIGALIVSLWGLYADRSRKKNLITYSILYGEIPICMASRSRSIGELFWRALIGIG
 VGSASPPTVSMGDMEVKRKRKVVAIISSAISISVGVLGMVGGFGKPGKGVWRFIA
 VSPNTIFAVVSIIVKVEKEKGAFERKEIGEYPKPLSDYAKVVKINL
 LFQFGTAGTIPGAIYFLVFFERRERGLSFEATLVFLVFGLGIVNGTLLGNGAS
 IVAKSRKFLPLICSIITLAQGIFTVTMIDYKMSLVLMLGFTASFTASITGPNTKF
 LLFLVFLVYFAKDVVERQKTMELAKNSQT"
 gene
 /db_xref="TM1022"
 complement("1341", 2102)
 /note="similar to GP:L77117 SP:Q591611 PID:1591465 percent identity: 55.70; identified by sequence similarity;
 putative"
 /codon_start=1
 /transl_table=11
 /product="transposase, putative"
 /protein_id="AA33603_1"
 /db_xref="GI:4981567"
 /translation="MPPRIRTYKLAICGHQNLQICEELNRIARINYKTMSLVQKIQH
 KKGWLSWITADKILRWNENIKHVHQFQVLYRQALKGFKAAKKDQAKPPTK
 KKYVLPFTKESAVSGLVQWVLSLGLGFGKFAGGWMSVAUGSLGALKVSYFWLICAV
 HLAIEVEETEKSANSVWMDIGVLRPTCDKPSKDEVKISVHGGULSSYRNRKLASL
 OSATRACKGCRRINKLPAKKRVRKRNQINDMHITTSSTGICLURKQGTTG
 DVTSGRERADYSIDANQKTHOWGPRKLEEMIRYKAEGPGIEVKLISANTSTCPVCG
 AKNKPNGRHYHKTCGFEYHRDGVGAINWKRPGTGVAGLAPVKGVRFRPHLCGG

Query Match 8.4%; score 89.2; DB 1; Length 20334; Best Local Similarity 54.3%; Pred. No. 6.7e-14;

gene
CDS
'gene'="TM1027"
'note'="similar to percent identity: 0.00; identified by sequence similarity; putative"
'codon_start'=1
'transl_table'=11
'product'="hypothetical protein"
'protein_id'="AAB36104.1"
'db_xref'='GI:491568'
'translation'="MSEEFGLGVQGNARYATATYNGTIVPQKAELIGFFSIGPPIHPLIKKVKVLEIPDIDFWWQNEPGSYGVDSVYFNGQECYRASCYENGLAELAEIKTVOTGLIVVQOLVE"
5922..6803
'gene'="TM028"
'db_xref'='TM028'
'note'="similar to GB:AE000782 percent identity: 60.48; identified by sequence similarity; putative"
'codon_start'=1
'transl_table'=11
'product'="ABC transporter, ATP-binding protein"
'protein_id'="AAD36105.1"
'db_xref'='GI:491569'
'translation'="MSVVERDTLYKGKSGRVGEGTFSVSEGEIFGFTGPNGAKTTIRKLGLIFPDSSLAETFGDKLFCKEITKNGVXIPGEVSFYPTVSEFLRSASDVYKDVYKELCKGVTSVLLDRKTRKBLISMGRKKVIAVQALMHRPLKLIIDEPINGLERVYKVRKGNEULEKLDAEVNRLKFENGTAEFLPSGNVERLIELRSKLSDLWEVERKVKVRRGENLEKPLKDAEVNRLKFENGTAEFLPSGNVERLIELRSKLSDLWEVEESLEELFMSYRREGEE"
6800..7591
'gene'="TM1029"
'db_xref'='TM1029'
'note'="similar to SP:P4233 percent identity: 50.27; identified by sequence similarity; putative"
'codon_start'=1
'transl_table'=11
'product'="ABC transporter, permease protein, putative"
'protein_id'="AAD36106.1"
'db_xref'='GI:491570'
'translation'="MNTWRDMMRYMKSTIAMIWVILQLQMLYAFPSMAKEETITKWIKVMPKAFKVLFGLEDMSIJMNLLAMWNSIYIVLGGFFASLGVRVSYREENEKTFPLSKRPSVSEWISKEKRSLSVHLIFDLSLFFSLMENANTISSPPEVEKFWFASQVEFHFTLMNPKVLSVLSLWVLSMISKTDAEFLKYFTPESSADPSVIRGGPPERSGLFFALVNLTALLSISVLESKKDLV"
7602..8204
'gene'="TM1030"
7602..8204
'gene'="TM1030"
CDS
'note'="similar to GB:S87512 SP:P43506 PID:623036 PID:247052 percent identity: 48.00; identified by sequence similarity; putative"
'codon_start'=1
'transl_table'=11
'product'="transcriptional regulator, TetR family"
'protein_id'="AAB36107.1"
'db_xref'='GI:491571'
'translation'="MUSKDAILKAAVEFGKKGYDRATDETAEKAVAKCLIFHEKNKELLYYQYMSYSTEKLQKEFNLKMRNRDFDEMERWEEKKLESASHEEADFLILVSYDEKLKRILLDEKQSVEFDEVRKULKDLOAEDTEETALKFLWFFSGFEVFLRTYQGKPEELLKRMTNTVEEVKVMRLTILKGMKTC"
complement(8229..8471)
'gene'="TM1031"
CDS
'note'="similar to GB:M60116 SP:P23171 PID:44907 percent identity: 73.61; identified by sequence similarity;
repeat_unit
repeated(8226..8471)
'note'="similar to GB:M60116 SP:P23171 PID:44907 percent identity: 73.61; identified by sequence similarity;
repeat_unit
repeated(8226..8471)

Matches 225; Conservative 0; Mismatches 183; Indels 6; Gaps 2;

Qy	427	GAACACCGCTCAGGTACTTGTCTACCCAAACACCCACGGAAACCTCTTTC	486
Db	15718	GGAGAGAGAGACGTTGTTCTATCGAACGCCACATCCACGGCCATGCTTGAA	15659
Qy	487	AGGGAAAGATGAGGAGTAGAACACGGGTTCTGCTTAATAGACCAAGCTAC	546
Db	15658	AGAGAGGAATAGAAAGAACCTCTGAAACGGGTTCTGCTGACGAAAGCTAC	15599
Qy	547	TATCATTACTCCGAGAACCTTCTGGAGACCCGCTCAACAGGGAGATACGGTAGTT	606
Db	15598	TACGATTCCAGGAACTTGTAGTATGGAATTCTGAAATACTGAAATACGAAATCGCTG	15539
Qy	607	TGTGAGACACTTCAAAATCGGTATGGATTCTGAACTGAGATACGGATTATAGGG	663
Db	15538	ATCAGGACTCTCGAAAGCTTTCCTGGAGCGCACGCTGGATACGGTAGGGC	15479
Qy	664	AAGGGAAATCTGCTCAGGTTAACAGGTGAGCTCCCTCAGGTGACTACCCC	723
Db	15478	TGGAGAAGTCTCATGACGCTACACAGGGTGGACCTCCCTCACTGAGCTAGTC	15419
Qy	724	TCTCAGGTGAGTGGCAAAGCTCTCAGGGGGAGAGATCTATGAAAGATA	783
Db	15418	TCCAGATGTGTTGCAAGGTGCTC---GATCACAGAGATCTTGAGAAGAC	15362
Qy	784	CAGGAGGTGTACAGAGCGGAAGGAGTACGACGAATAGGAAATAGAA	837
Db	15361	AAGTCATCGGGAGAACGGAGAGATGAGTGCCTCAGGGAAATGGGA	15308

RESULT 9
NMA622491
LOCUS NMA622491 326301 bp DNA linear BCT 04-DSC-2000
DEFINITION Neisseria meningitidis serogroup A strain z2491 complete genome;
VERSION AL162757.2
KEYWORDS segment 6/7; GI:7380371
SOURCE
ORGANISM Neisseria meningitidis z2491.
REFERENCE
AUTHORS Neisseria meningitidis z2491.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
1. (bases 1 to 326301)
Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klein,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
Holroyd,S., Jagels,K., Leather,S., Moulton,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis z2491.
Nature 404 (6777), 502-506 (2000)

TITLE JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT Notes:
Details of *N. meningitidis* sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
source
1. 326301
organism=Neisseria meningitidis z2491,
strain="Z2491"
db_xref="taxon:122587"
notes="serogroup: A"
note="REP 2; humfs hit to HMM REP 2 (1 - 109), score:
196..311

VERSION	AE000963.1	GI:	2689286
KEYWORDS	Archaeoglobus fulgidus.		
SOURCE	Archaeoglobus fulgidus		
ORGANISM	Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales; Archaeoglobaceae; Archaeoglobus		
REFERENCE	1 (bases 1 to 22014)		
AUTHORS	Kleck, H.P., Clayton, R.A., Tomb, J., White, O., Nelson, K.E., Richardson, D.L., Kerlavage, A.R., Graham, D.E., Kyrides, N.C., Richardson, D.L., Kerlavage, A.R., Graham, D.E., Kyrides, N.C., Fleischmann, R.D., Quackenbush, J., Lee, N.H., Sutton, G.G., Gill, S., Kirkness, E.F., Dougherty, B.A., McKenney, K., Adams, M.D., Loftus, B., Peterson, S., Reich, C.J., McNeil, L.K., Badger, J.H., Glodek, A., Zhou, L., Overbeek, R., Gocayne, J.D., Weilman, J.F., McDonald, L., Utterback, T., Cottrell, M.D., Spriggs, T., Artiach, P., Kaine, B.P., Sykes, S.M., Sadow, P.W., D'Andrea, K.P., Bowman, C., Fujii, C., Garland, S.A., Mason, T.M., Olsen, G.J., Fraser, C.M., Smith, H.O., Woese, C.R. and Venter, J.C.		
TITLE	The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon <i>Archaeoglobus fulgidus</i>		
JOURNAL	Nature 390 (6658), 364-370 (1997)		
PUBMED	9389475		
REFERENCE	2 (bases 1 to 22014)		
AUTHORS	Klein, H.P., Clayton, R.A., Tomb, J.-F., White, O., Nelson, K.E., Richardson, D.L., Kerlavage, A.R., Graham, D.E., Kyrides, N.C., Fleischmann, R.D., Quackenbush, J., Lee, N.H., Sutton, G.G., Gill, S., Kirkness, E.F., Dougherty, B.A., McKenney, K., Adams, M.D., Loftus, B., Peterson, S., Reich, C.I., McNeil, L.K., Badger, J.H., Glodek, A., Zhou, L., Overbeek, R., Gocayne, J.D., Weilman, J.F., McDonald, L., Utterback, T., Cottrell, M.D., Spriggs, T., Artiach, P., Kaine, B.P., Sykes, S.M., Sadow, P.W., D'Andrea, K.P., Bowman, C., Fujii, C., Garland, S.A., Mason, T.M., Olsen, G.J., Fraser, C.M., Smith, H.O., Woese, C.R. and Venter, J.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-DEC-1997) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA		
REMARK	In order to show the genes in ascending order on the genome, the origin of this version has been moved by TIGR to position 203570 of the original version and the opposite strand is shown from the original version.		
COMMENT	On Dec 16, 1997, this sequence version replaced gi:2648504.		
FEATURES	location/qualifiers		
source	1. 22014 /organism="Archaeoglobus .fulgidus" /db_xref="taxon:2234" /gene="AF2014" 61. 1266 /gene="AF2014" /product="sugar transporter, putative" /protein_id="AAB8936.1" /db_xref="GI:2048517" /translation="MCEFSKEKLVLLAAGWSHDAVGILNLFLAEPIMREFNUGTEIGFIFSAQICIFCAGMLGMDGRFRKLNLISIIVWFSVSLAPFWTLTFLWIAEFAHAIWVMSWPTIQRQYRERASIVLSPVLSVLSFLVPIGFPSHAKFGVNSSEMFPKRATASISGLGCLMAFTYPLVQFSLYQGMLGGIATFLMCIPWNLPETRGRELPE"		
CDS	gene gene="AF2015" 1276. 2907 /gene="AF2015" /note="similar to GB:17717 SP:057725 PID:1591003 percent identity: 34.07; identified by sequence similarity; putative"		
RESULT	10		
ACCESSION	AE000963		
LOCUS	Archaeoglobus fulgidus 22014 bp DNA linear BCT 15-DEC-1997		
DEFINITION	Archaeoglobus fulgidus section 144 of 172 of the complete genome.		

gene
CDS
/product="acetoacetate synthase, large subunit (ilvB-3)"
/db_xref="GI:2048516"
/translation="MOGAEVAKALLREKIDCAFLSPSSEILPVCEYLEAEGVRIIFT
RHEOAAGNAADAGYARVTRPCCFIVPTGCGLNLMPALMQAYASSWVAIAGHSRYV
NLDRAEAEEDIGLAWKEYIWKWSRSLPFPFTKHLHEVQEARRAISGRGPVLIBPKD
VLNADIDGKTYELAERNYRTGTRVCEETYIYRALEMUNAKEPLIVAGSYWWSRSE
NFMIAFAEKISIPVSGLGGCTSPDHPLFAGOSASAPWQSDFLVNGTREFEL
GFGQKFKVWHVIDACELAKRNVDIAITSCDAGYFFRKALANLEGEKFENWARPI
MEGVNRMFEELAKGKEDPMFQLRNLSDRIVLQDGTTTANGLLYIKKGKV
ISSOOPFGHIGAGTINGIAAKAVPKRKEWVTDGSELFNGEIDTAVRHGQWVV
VNDRSGWGLVNHTRLISTKSARASIGVMNPAAKYDRIFAESELBEAGYELVEKSGEVKE
AVKRAFDSDLALFADVRLVKLDEISPLAYLLSGSG"
/gene= "AF2016"
2909. . 3832
/gene= "AF2016"
/product="UDP-glucose 4-epimerase (gale-2)"
/protein_id="AAB8934.1"
/db_xref="GI:2048515"
/translation="MAVLITGGAGLFLGRFVAKHFDSPVILDKLEGKGGIFERGSVTAW
SDIAEVFRKHEIEGVTHAAELSTYKSHVDAFRANREGALNILEACRVFVEKVF
TSSVSVGGRSYPTESYRDPITFYGATRACCSLIGITYSYTIGDIDEAVRREPILVG
PFRRMGASVAFSSLDDAFFFNGJSVITPEDETRLPVXVKDAELVERLYSIDRVS
SRIEFLDWLRWKSAEIEVEDHLKLAEEV"
gene
CDS
/product="3-hydroxyacyl-CoA dehydrogenase (hbd-9)"
/gene= "AF2017"
3829. . 4674
/gene= "AF2017"
/note="similar to GB: M31799 PID: 144733 percent identity: 35.36; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="3-hydroxyacyl-CoA dehydrogenase (hbd-9)"
/db_xref="GI:2048514"
/translation="MKVTFVAGLMGRGIALATASKHWWLQDSEKALEAROITE
EUSKIEFTTTELEKDDIYMERAVEFDTINTKVEYLREVERETNAPCNSNTVSD
IAERLDSFPRFLGHWMMPHPWAPLVELVTSRSTDTSKVAPEGFLBLGKNNVCKD
QSNVNRFNAVLSSEASMRMEEGTRAEDYDWRWKHGLYLTFGPGLNLDYGLDWA
YASLYVYRFGDEKEKPPFWLQKIKKEFVGKAGKGYEYOPKAYEERVERKLKLIR
FIGLE"
4671. . 5405
/gene= "AF2018"
/gene= "AF2018"
/note="similar to GB: 54844 percent identity: 26.98; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/db_xref="GI:2048513"
/translation="AINEATIVAMFLTLIVAFFVWVTFGGSAIILSPISNIMPKEA
VVILLILSFVNVFLVFRTRKERMNSLRELYLGFLAAGTIFLFGFLPOETVGLTIGVA
MGMAVMLLGNVFRTRKERMNSLRELYLGFLAAGTIFLFGFLPOETVGLTIGVA
RSIMITYLWVNTATLSAFAFGHVNQQLTINFLILAPSIIAYALGSRLSRMDGER
LKKRILSAAVVLNLSLWVILRYLGGLIG"
5422. . 6537
/gene= "AF2019"
5422. . 6537
/gene= "AF2019"
/note="similar to SP: P45513 PID: 493086 percent identity: 35.71; identified by sequence similarity; putative"
/codon_start=1

gene
CDS
Query Match Best Local Similarity 8.1%; Score 86; DB 1; Length 22014; Matches 476; Conservative 47.4%; Pred. No. 5; 5e-13; Gaps 6; Gaps 476;保守性 47.4%; 预测数 5; 比较数 476; 缺口数 6;

Qy	61	GTGAGCCCTTCCTAACGAATTCCCCTACGACTTCCCCGAGGAGTAACACAAGGGCC	120
Db	9660	GTCAACCTTCTCACAAAGAGATCCTATAAGCCGAGGAGTTAGGGCTAC	9659
Qy	121	TTAGAAGATTTAAAGGTTCTGACAAATACCGACGCCAGGAAAGAGTGAA	180
Db	9660	CTTAGGCTCTGAAATCATTC-----AAGATATCTAACGGACTAGGCCGCTAC	9713
Qy	181	AAAGCGGTTCTGGATTTCGGCTTAGGAGAAATTAGTTCTCGGTAACTG	240
Db	9714	AAAAGCCCTGCGAGATACAGGGCTTGAAGTGAACATAGCTGGGAGTGGAA	9773
Qy	241	TGGAGAACCTATGACTCTCAATAGCTATAGCTTACATACCGGTTAC	300
Db	9774	GCATCTGAGTAAATCTCTGCTCAAGCTCTGCTGGAGAGCTGGACGGGTGTC	9833
Qy	301	ATACCTGTTCCACCTTCTGAGAGATAAGTGGAAAGTCTCGGAAAGACCCCTC	360

	REFERENCE
Db	9834 ATTCCAAATGCCCTCTACACCCCTTACTTCACTATGCAATGTTGGGAGGCCATCCATC 9893
Qy	361 GTAAAGGTCACTGGAGAAACCTTCATAGACTTAGAAGAAGTTAGATAATA 420
Db	9894 AGCTTCCCCTTTTGAGGATACAGGTGA-----TCCAGTCATC 9938
Qy	421 GAGAAGAAAACCGTCTCGGTACTTTGCTTACCCAAACACCCACGGAAACCTC 480
Db	9939 GCCGAGAAAGGGAAAGCTCTGTTTCTCGCTCTCCACACCCACGGAAACAGC 9938
Qy	481 TTTCAGGGAAAGATGGAGATAGAACAGGGTTCTGGTATAGAGAA 540
Db	9999 CTGAGCAGAAACTGTCAGAAGAGTGGCTGAGTCAGGTTAGGTACACAACTG 10058
Qy	541 GCCTACTATCATCTACTCCGGAGAACCTTCTGGAGAGCGCTCAAAGGGATACG 600
Qy	6059 GCCATGTGGATTCAGCATGTTCAAGATGATGAGTGGTTAGGTTACACAACTG 10118
Db	10179 GTCTSGAGTGGAGATTGCTGAGGCAATTGGGTAGGGATTTA 657
Db	10119 ATGGTGTGAGAACGCTTTCAGTTTCGGGTTGCGGAATGAGGTGTTATGCT 10178
Qy	658 ATAGGAAGGGAAATGGCTCAACTAACAGGTGAGCTCCCTCAAGGTGAGCC 717
Db	10239 TATCCGAGCTGAGGAACTTCAAA---ATCGGTAGGGAGTTAGGGTAAAGGTTA 10238
Qy	718 TACCCCTCTAGGATGGCAAACTCTCTCACGGGGAGGAGGGCTATGAAACTGAA 777
Db	10239 TATCCGAGCTGAGGAACTTCAAA---ATCGGTAGGGTAAAGGTTA 10237
Qy	778 AAGATACAGGGATGTAACAGAGGAGATGAGCAATGAGAAATAGAA 837
Db	10298 AAGGGACAGA - -TGTCAGGGAGGGAGGAGGGCTATGAAACTGAAATGAG 10355
Qy	838 GGATGAGGTTCGGAGTAAAGCTTAACCTCTCTGCTTAAGGGACAAAGGGGC 897
Db	10356 TGGCTGAGTTACCGTCAGGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 10415
Qy	898 CACGAGTTATCAGAGCTACTGAAASGATGCTGCTGCTGCTGCTGCTGCTG 957
Db	10416 GTGGTAG-----AGAGCTGGCGGAAGGGTTATGTGAGGGATGCAAGTGT 10469
Qy	958 GAAAGACTCCAAGTGCTCAGGTAAGCGTAAAGGAAACAGGT 1017
Db	10470 GGCTTGAGGGTCTGTATTCGCAATACAGTGGTAAGGGATAATGACAGGTG 10529
Qy	1018 CTGGAACACTGGAGGAGGATATAATCCCTTCAAGCTCT 1061
Db	10530 ATTGAGGCCCTAAGAGAAATTAGAAGAGGTCTCGCCTTAT 10573
RESULT	11
AL646072/c	
LOCUS	AL646072
DEFINITION	188050 bp DNA linear BCT 07-DEC-2001
SEGMENT	Segment 16/19.
ACCESSION	AL646072 AL646052
VERSION	AL646072.1 GI:17429824
KEYWORDS	Ralstonia solanacearum.
SOURCE	Ralstonia solanacearum.
ORGANISM	Ralstonia solanacearum.
REFERENCE	(bases 1 to 188050)
AUTHORS	Arlat, M., Billaud, A., Brotté, P., Camus, J.C., Catolico, L., Chandler, M., Choisne, N., Claude, I., Renard, C., Cuninac, S., Demange, N., Gaspin, C., Lavie, M., Moisan, A., Robert, C., Saurin, W., Schilex, T., Siguer, P., Thebault, P., Whalen, M., Wincker, P., Levy, M., Weissenbach, J. and Boucher, C.A.
TITLE	Genome sequence of the plant pathogen Ralstonia solanacearum
JOURNAL	Unpublished
COMMENT	
FEATURES	
source	1. 188050
gene	/organism="Ralstonia solanacearum" /strain="RMI1000" /db_xref="taxon:305" /transl_table=11 /product="PROBABLE TRANSMEMBRANE PROTEIN" /protein_id="CAD1659_1" /job_xref="GI:17429825" /translation="MIALLSMASVCGNSFAKTLFPAKGAGTYRITIGAILLA LWRPWRLRHDRDAGTALGYTLASMNLFLYSLTRLPGIAIEFVGPVLAVAL SRALDEVVIGAVAGLIGVAGLIGVAGLIGVAGLIGVAGLIGVAGLIGVAGL PAGQATPSLGMVAGLAIPEFGVVOAGPALAAPSILVAGGLGVLSVASSVYSL HLSGRFTSVLISLEPAIGALAGAVVHLHEHLRSARQWAVIAVITAASAGCATARSRAA ER"
CDS	complement(R36..2024) /gene="RSC2803" /note="RSC0302" /complement(R36..2024) /gene="RSC2803" /function="miscellaneous; hypothetical/partial homology" /note="Product confidence : hypothetical Gene name confidence : hypothetical predicted by Codon_usage predicted by FrameD predicted by Homology predicted by FrameD /evidence=not_experimental /transl_table=11 /product="HYPOTHETICAL PROTEIN" /protein_id="CAD1651_0_1" /db_xref="GI:17429826" /translation="MAEATIESFAQYDQARARFLDARARGLATERAVPHHALGPAG EPISIPACFLDFDRPLWVLSINGVEPGCGSGQVGVLJLDAFLFARLAAGRFL LYHAVNYGFAHRYRNEDVNDLNRSADFAVASANPYLEVDPILLPWTWPPAAN QALQQYLAMHGSEALRDMTQYALPDMYGGATCWSTAQACILRQHADAPR LAWIDHNGLGANGHGERIFSCADPRLERATIGTWAADVRLIAAGSVSSVVEGLVD RAGLPAEKVYIITLEFGTLPWMAVNOALRADHWLRHGTSPDQAATRKGLDFA YCPDANKMVIAQARYAVLQLQARFFSG" /gene="RdB" /note="RSC2804; RS00301" complement(R124..3326) /gene="RdB" /EC_number="1.17.4.1" /function="small molecule metabolism; central intermediary

ORIGIN		Query Match Best Local Similarity Matches	Match Score 6.8%; Pred. No. 3.4e-99; 0; Mismatches 563;	DB 1; Length 1480;
DDD	167893 CGCTCCCGCCGGCTAGGCCCTGGGTGTCAGAACGGACTCGCTTACGGCT 167834	Qy 545 ACTATCATTACTCGGAGAAACCTTCTGGAGACGGCTAAAGGAAGATACTGGTAG 604	Qy 2 TGATACCCGAGGTTAAGGAACCTTAAGAACGAGGACTCCCGCTTACGGCT 61	4;
Db	167833 ATAGCGGTTGCCACACACTGGATGGGTTTATAGGGA 664	Db 238 TGCCTGGCCCGAGCTTGCTGGGAGCAAGGCTTATGGGGAGACGGTTCTCA 297		
Qy	605 TTGTGAGGACACTTCAAATCGGTATGGGTTAAGGTAGGTTTATAGGGA 664	Db 238 TGCCTGGCCCGAGCTTGCTGGGAGCAAGGCTTATGGGGAGACGGTTCTCA 297		
Db	167773 TGTGCGCACCTTCAAGCTGGCTGGGGTACGTGGCGGCC 167714	Qy 62 TCGGTTTCTCTACGATTCCCTACGACTCCGAGATAACGGAGATAAAACAAAGGGCT 121		
Qy	665 AGGGGAAATCTGCTCTCAGAAATTAACTGGTCAAGCTGCCCTAACGTTGACCT 724	Db 298 TCAAGCTTGATGCCATGGGAAACCCCTATGSCCTTACCGGGCTTGCCTAGCA 357		
Db	167713 ACCCGAAGTGGATGCCGAGCTGACAAGGTCGCCGCGCTAACGTTGACGTCTGA 167654	Qy 122 TAGAAGAATTAAAAAGTTCCTTGAAACAAATACCGAGCCCCAGGGAAAGGTTAA 181		
Qy	725 CTAGGTTATGGAAAAGTCTCTACGGGGAGAGAAATTCTPAATGAAAAGATAC 784	Db 358 TGGAGGAGCTGGCCAGCGCCCTCATCGCTATCCGAGCACAACCCGACCCCTCA 417		
Db	167653 CCGAGGACCCGGCTTCGCGCTTCTACGACCTGGGAGTSCCAGCG 167597	Qy 182 AAGCGGGTCTTCGGATTTCGGCT - - TAAGAAGAAATTTAGTCTCGTAACG 238		
Qy	785 AGGAGGTTGTAACAGAGGGAGTACGACAAATAGAAGGGATGT 844	Db 418 TGGAGGAGCTGGCCACATGGCTGGCTGGGAATGAACTCATCGTGGTAACG 477		
Db	167596 CGACCTCGTGGGGAGGTACCGGCTGATGGCTGTCAGGGCTGA 167537	Qy 239 GTCGGGGAGAACTCATACATCTCAATAGCTATAGTGTAACTTACATCCGGTT 298		
Qy	845 AGGTTCCTTCGAGCTTAACCTCTGCTTCAGAACGCCCTAACCGCCACAGGG 904	Db 478 GTCGGATGAGGTGATCAGATTCTGTTGGACGGGGACCCGATCA 537		
Db	167536 CGGTGTTCCGACGCCGCCCCAAATTCGCTGGCGATGCCG - - CGCC 167480	Qy 299 ACATACCTGTTCCACCTTCCATGTAAGGATAGTGGAAAGTCTCGGAAGACCCC 358		
Qy	905 TTATCAGGACTACTGAAAAGGATGTCGCTGGGAAACCTATCTACATGAAAGAC 964	Db 538 TGGGGTAGACCCCGAGTTCAGCTGGGACTCTGCTGGGAGCTGCTGTC 597		
Db	167479 TGTTCAGGGCTTGTCTTCAACGGGCGCTCTCAAACCTCAGTAAAGCACCGT 167420	Qy 359 TGTAAAGGTTCAACTGGGAAANACTTGTATAGACTTAAAGAATTTGAAATTAA 418		
Qy	965 TCCAA -- AAGTGGCTCAGGGTAAAGGTAGGGAAACACAGTTCTGG 1021	Db 598 TGTGGGCTATTCCCTGTATGGGACTCTGCGCTGACCTTGCTGTC 657		
Db	167419 TGGTGGCCAACCTGCTGGCGTGTGACGGTCAACCCGAAAGAACCGTCTTG 167360	Qy 419 TAGAAAGAAAACCCGTTCTGGGTACTTTGCTTACCCARACAAACCCACGGGAAAC 478		
Qy	1022 AACGACTGG 1030	Db 658 TCCCCGCCAACACCCGCCATTTTGTGACTGGCCACAAATCAGTGGAGC 717		
Db	167359 ATGCCTTGT 167351	Qy 479 TCTTTTCCAGGGAAAGTGGGAGATAAGAAC - - AGGGTGTCTCTGTAAATG 535		
Qy		Db 718 TATTCGGAGACGATCTGGGGCATTTGTGGGTGTGGTG 777		
Db		Qy 536 ACGAAGCCATACATCATCTCGGAGAAACCTTCTGGAAAGCGCTCAAAGGGAG 595		
Qy		Db 778 ATGGGGCTATTAGCCPTCAGTCAGGGACCTTGTGGATCACCTGGACCCCCA 837		
Db		Qy 596 ATACGGTAGTTGGAGGAACCTTCAAAATCGTATGGGAGTTAAGGTAGGGATT 655		
Qy		Db 838 ACCCTCCCTGTGCGGACACATGCCAGGAGCTCGGGGATGGCTGGGATG 897		
Db		Qy 656 TAATAGGGAGGGAAATCTGCTCAGAAATTAAAGGACTCCCTCAACCTGTA 715		
Qy		Db 898 TGGGGGGCCGGCGCTGGGATCTGGGACTCTGCGCTTACAATATCA 957		
REFERENCE	Acidithiobacillus ferrooxidans	Qy 716 CCTACCCCTCTCAGGTGATGGCAAAGTCTCTCAGGGATCTCTTAATGG 775		
AUTHORS	Bacteria: Proteobacteria; gamma subdivision: Acidithiobacillus.	Db 958 ATGTACTACCCAGGGACTGCTCTACCT -- GCCTACACCCGAAAGTGTGAATG 1014		
TITLE	1 (bases 1 to 1480)	Qy 776 AAAAGATAAGGGTTAACAGAGGAAAGGATCTACGAGAAATGAAGRAAATAG 835		
JOURNAL	Acidithiobacillus ferrooxidans ferrooxidans	Db 1015 CCCAGGCCAAATTCTGGTGTCAAGGGAAACGGCTCTCAAGCCAT -- CGCTGCT 1071		
PUBLISHED	gene, partial sequence.			
REFERENCE	Acidithiobacillus ferrooxidans			
AUTHORS	Acidithiobacillus ferrooxidans			
VERSION	Bacteria			
DEFINITION	1 (bases 1 to 1480)			
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS				
DEFINITION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
REFERENCE				
AUTHORS				
LOCUS</td				

ILGYAILENGKDVVPLEGRNTISDFNIGEKIRSHTPPPHTEAAKEIKAEMSQAERFEFKW
EEFACIKCYACROVCPMCYCRRCIVDVOQONSTSHTLIGNEFWNLVRAPHLAGC
AAGACGACDRACPVNTPRLINYRMCKEVRSADFIVAGENSQKFVLASFQDDPFTIFL
gene 6005 . 7018
/gene="hydG-2"
/note="CT1249"
6005 . 7018
CDS
/gene="hydG-2"
/codon_start=1
/transl_table=11
/product="hydrogenase/sulfur reductase, beta subunit"
/protein_id="AAM72479.1"
/db_xref="GI: 2164725"
/translation="MTRILLNKLDCLIAWOKAGESVYLAPVVKRHEMSCEGYOKSGD
NALDLVTERITKODFPOPEPLKRYRKQIDSEIMTPPEKKRVEPGVPCDAASGL
FIDDPFLGWDXKDDDWERREKSVTIVCAQADEFCMTSVSKLSPDSTKGADMLRP
ISDGSGWOWEAWSDEGDTALVTSSLLQESSAEAAPPOVAEKFEDVKEVMENLADKEN
FESQEWKDIARVCVGGSCTFLCPICHCFIJQDGGDTIQGIRRKKNWDSCSPELFMTHT
SGHNPNTQSTWRORIMHKFNYRGKFGYNNSCSGCRTRCPVDMGITTELQLAITN
LPR"
gene 7045 . 7881
/gene="hydG-2"
/note="CT1250"
705 . 7881
CDS
/gene="hydG-2"
/note="identified by match to PFAM protein family HMM
PF00175"
/codon_start=1
/transl_table=11
/product="hydrogenase/sulfur reductase, gamma subunit"
/protein_id="AAM72480.1"
/db_xref="GI: 21647226"
/translation="MYSPPFMRVSYKRAAPGNTNLKLEVKQGDHEFFKANVRTGM
FPJGFEGGSTPCVASPERTEKTFECTRQSGRTVSTLNDAGDIVTRFGYGNR
DEWKORDVRLVLTDPGGEPDWDHVGPVTLQEAESPENTAVLCGPPMIKF
TLLALEKLGFRAENVTTLERNMKCGIGKRCNVGSIYIKEGPVFTAEVQAMPQA
DL"
7887 . 8000
/gene="CT1251"
/note="hypothetical"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM72481.1"
/db_xref="GI: 21647227"
/translation="MRMDTKRGRKRSUFLCKSFLIHDFPHPNILLFRHH"
gene 8028 . 8591
/gene="pth"
/note="hypothetical"
/product="hypothetical protein"
/protein_id="AAM72482.1"
/db_xref="GI: 21647228"
/codon_start=1
/transl_table=11
/product="peptidyl-tRNA hydrolase"
/protein_id="AAM72482.1"
/db_xref="GI: 21647228"

Query Match 6.58; Score 68.8; DB 1; Length 11265;
Best Local Similarity 45.4%; Pred. No. 3.8e-08;
Matches 447; Conservative 0; Mismatches 522; Indels 15; Gaps 5;

gene 61 GTCAAGCTTTCCTCAAGAAATTCCCTAACGACTTCCGAGGAGATAAAAAGGGCC 120
/note="peptidyl-tRNA hydrolase"
/protein_id="AAM72482.1"
/db_xref="GI: 21647228"

Db 11010 CTCGACCAAGTTGCCATGGCCATGGGATATTCTGGTACCGGGC 10951
Qy 181 RAAGCGGTCTTCGGGATTTCGGGTTAAGGAAGAAATTAGTAGTCTCGGTAACGGT 240
Db 10950 ATGGCGGCCTATGCGATCTTTCGGGTTAAGCGGAGTTGGTATCATGCAATGGT 10891
Qy 241 TCGGACGAACTCATATACTACCTCTCAATAGCTTACATACCGTTAAC 300
Db 10890 TCGAACGAGATGCGCTAACCCATCTGGGCTGGCTGGCAAGGTGCTT 10831
Qy 301 ATACCTGTTCCACCTTCGGGATAGTGGAAAGTTCTCGGAGACCCCTC 360
Db 10830 ATTCCAGACGCCCTCGGTTTGCCTTACGAGCTTCAGCAGGCTGGGGT 10771
Qy 361 GTAAAGGTCAACTGGACAAACTTGTATAGCTAGATAAGAAGTAAATAATA 420
Db 10770 GTCPAGGTTCCGATGTCATGACGATCTCTCTTGACGTCATGCGATCATCGAAGCCGC 10711
Qy 421 GAGAAGAAAACCGGTTCTGGGTACTTTGTCTTACCCAACCCCCACGGGAAACCTC 480
Db 10710 AGGCCGGAAGGGCAGATTATGGCTCTCGGCAACATCCGACCAACTCG 10651
Qy 481 TTTCAGGGAAAGATTCAGGAGATAAGAAACAGGGTTCTGNG --TAATAGAC 537
Db 10650 CTGTCCTCACGAGGATTCACCGGATTCGCAAGCTGCGACCCCATGTGCGTGGCTGAT 10591
Qy 538 GAAGCTCATATCATTTACTCGGAAACCTCTGGAGACGCCCTAACAGGAAGAT 597
Db 10590 GAAGGGTAGTGTGAGTTTCGGTGGCAATCGCGCTGACCTGACCCCTACCGC 10531
Qy 598 A---CGGTAGTTGGAGACATTCAA--AATCGTATGCGAGTTAACGGTAGGG 651
Db 10530 AACCTGATCTCTGTCGCGACCACTTCGCGCTTCCANGGGAAATGCGTATGGC 10471
Qy 652 ATTAAATAGGGCAAATGCTCTAGAAATTAACTAACAGGTGAGACTCCCTCAC 711
Db 10470 TTGGGATGSCCAATCGGAGCTCTGGGAATCTCCAGCGGAATTTCCCTTGGC 10411
Qy 712 GTGACCTACCCCTCTCAGGTGATGCCAAAGTCTCCACGGGAGAAATTCCCA 771
Db 10410 TCGAGCCGACTGGGAAATCACTCTCATGGCTGTGCTG --GAAAACACTACTCTG 10354
Qy 772 ATGGAAAAGTACAGGAGTTGTAACAGGCGGAAAGATGTGACCCCTAAC 831
Db 10353 ACCGAGCCGCTGCACTACCTGCGCGATCGGGGAGCTGACAGACA 10294
Qy 832 CCCGCCACGAGGTTTATCAGGATCTCGTACAAAGGGATCTCTGCGTAC 951
Db 10233 GCATC --CGGGTGTTCAGAAGTCAAAATGGGGTGGTGGAAAGTGTG 10177
Qy 952 TACATGGAAAGACTCCAAGGTGCTCAGGGTAAGCTGGAAACGGAAACAC 1011
Db 10176 GGATACCCGCTTATGGAGAACTGCGCTCCGCTCATGTTGGTCTCAGGGAGGAAACGAC 10117
Qy 1012 AAGTTCGGAAAGACTGGCTCAGGGTAAGCTGGAAACGGAAACAC 1035
Db 10116 CGCCGTTGGAAACTGCGCTCCGCTCATGTTGGTCTCAGGGAGGAAACGAC 10093

RESULT 14
AV117255 LOCUS Arabidopsis thaliana putative histidinol-phosphate aminotransferase
DEFINITION (At5g10330) mRNA, complete cds.
ACCESSION AV117255
VERSION 1
KEYWORDS FLI_CDNA.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytina; Magnoliophyta; eu dicots; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1285) AUTHORS Yamada, K., Banh,J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C., Toriumi, M., Wu, H.C., Yamanura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE JOURNAL

Arabidopsis Open Reading Frame (ORF) Clones
Unpublished 2 (bases 1 to 1285)
Yamada, K., Banh,J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C., Toriumi, M., Wu, H.C., Yamanura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.

Submitted (03-JUN-2002). Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEc (SSP) Consortium members constructed and
sequenced the pun1 (ORF) clones using the RAFL cDNAs: Yamada, K.,
Banh,J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,
Toriumi, M., Wu, H.C., Yamanura, Y., Yu, G., Bowser, L., Chen, H.,
Cheuk, R., Jones, T., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J.,
Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A.,
Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEc) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP
/PGEc) contributed equally to this work as PIs.

FEATURES SOURCE

Location/Qualifiers 1..1285
/organism="Arabidopsis thaliana"
/db_xref="Taxon:3702"
/chromosome="5"
/clone="U11135"
/note="This clone is in pun1_51.
ecotype: Columbia"

gene 1..1285
/gene="At5g10330"
/gene="At5g10330"
/codon_start=1

/evidence="experimental
/product="putative histidinol-phosphate aminotransferase"

/protein_id="AA51330_1"
/db_xref="GT: 21436321"

/translation="MGIVINQSPSPSTHSSESNLRKSRALKKPFCSIRNRYCAQSS

SAAVDESKNTIMGDSEFLRPHLROLAYQPILPEFEVLSAQGRKPKDIDVLDANEINPYG

PPPEVEAAGMKKEPKVYRDPQRPLRDAQDSCLPESLTLVGGADEJLDLMRCV

LDDGEKFLIDCPPTFSMVFDAVNGVTKPRNDFSLNDRDAEVVELEKPKCIFL

TSPNNPPGSLLISSEDDLKLLEMPYLVLVDAYIETSGVSEMRKVKYENLIVRTFS

KRAGLAGLRLVCGYAFPLSILEYNNPNSVAGEVAAALNSKGKLEVEDVDAV

RERERLFLGLKKEVPFLNPYPSNFLCEVTSMDAKLKREDLARMGVNVRHNSQEL

KGYVRVISAQKPEHTDYLMECLQFY"

1255..1285
/gene="At5g10330"

3' UTR

358 a 241 c 302 g 384 t

BASE COUNT ORIGIN

Query Match Score 64.2; DB 8; Length 1285;
Best Local Similarity 48.3%; Pred. No. 6 2e-07;
Matches 180; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

Qy 190 CTGGGATTTCGGCTTAAGGAACAAATTAGTTCTCGTAAACGGFTCGGAGCAA 249
Db 394 CTGGCTCAAGACTCTGGTTGGAGTCATACTGTAGGCTGTGAGTGCCT 453

Qy 250 CTGATATACTACCTCTCAATAAGCTATAGGTGAACTTACATACCCGTTACATACCTGTT 309
Db 454 CCTATGATTGATCAGAGAAGTTATAGATTCCTGGGA 513

Qy 310 CCCACCTTCCATGATCAGATAAGTGCAGAACCCCTGTAAAGGTT 369
Db 514 CCACACTTCATGATGATGTTGCTGGCTGTGAATGGCAGCTTGCTATTAAAGTT 573

Qy 370 CAACTGGACGAAAACCTTGATATGACTTAGAAGAGATTGAAATTAGAGAAAGAA 429
Db 574 CCAAGAAACCCGGATTCAGCTTCAAGTGTAGACCGCATCGCTGAACTGAGA 633

Qy 430 AAACCGGPTCTGGTACTTGTCTTACCCAAACAGGGAAACCTTCTTCAGTGGGTATCATCGGAG 489
Db 634 AAACCCAATGCAATTCCTAACTTCACATCCAGTGGGTATCATCGGAG 693

Qy 490 GGAAAGATTGAGGATAAGAACAGGGTTCTGTAAATAGACGAAAGCCTACTAT 549
Db 694 GAGCATCTGTGAAGATTCTAGAATGCTGTTCTAGATGAGCTTAACATC 753

Qy 550 CATTACNCGGAG 562
Db 754 GAATTCCTCAGGAG 766

RESULT 15 LOCUS AY050832 DEFINITION Arabidopsis thaliana putative histidinol-phosphate aminotransferase (At5g10330) mRNA, complete cds. ACCESSION AY050832 VERSION AY050832.1 KEYWORDS FLJ_CDNA. SOURCE Organism Arabidopsis thaliana. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; Rosidae; eurosids II; Brassicales; Arabidopsis. REFERENCES 1 (bases 1 to 1570) AUTHORS Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. Unpublished

JOURNAL Unpublished

Yamada, K., Liu, S.X., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. Direct Submission

Submitted (08-AUG-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGC (SSP) Consortium members carried out the sequencing and annotation of the R AFL cDNA: Yamada, K., Liu, S.X., Phan, P.K., Banh, J., Banno, F., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin, Neumann, G., Kim, C., Kozen, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGC)

Contributed equally to this work as PIs.

Location/Qualifiers

1..1570

/organism="Arabidopsis thaliana"

/db_xref="Taxon:3702"

/chromosome="5"

/clone="R AFL 08-15-J07 (R11135)"

/note="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XbaI insert.

ecotype: Columbia"

1..1570

/gene="At5g10330"

1..133

/gene="At5g10330"

1..134

/gene="At5g10330"

/db_xref="GI:1529219"

/translation="MVGIVNQGSPSFHSSESNLRSKRSALKKPFCSTIRNRYCAQS

SAAVDESKNNTMGDSFIRHLRQLQAYOPLPPVLSAQRLPEPDIVKLDDANEENPYG

PPPEVEALGMKEPVYDPPQSRRLDADQDSGLESEBYIYVGCGADELLDLMRCV

LDPGENKLIDOPTSMYVEDAVAGVQVLPNPKDFSLNVDIAEVVLEPKCIPL

TSPNNPDGSTSIEDDLKMPFLVLDPEVWKKYENLIVLRLTES

KRAGLAGLRYVGAPPLSLIEYLWRAKOPYNVSAGEYTAALAALSNCRYLEYVRDALY

RERERFLGLIKEKVPLNPFTSYSNFILCEVTSGMDAKKLKEDLAKMGYMWRYHNSQEL

KGYVYVSAKGPHETHTDYLMECLQFY"

1..388

/gene="At5g10330"

1..1553

/gene="At5g10330"

/note="compared to genomic sequence"

/replace="9"

3' UTR

/misc_difference="At5g10330"

BASE COUNT

444 a 298 c 344 g 484 t

ORIGIN

Query Match 6.0% Score 64.2; DB 8; Length 1570;

Best Local Similarity 48.3%; Pred. No. 6.4e-07;

Matches 180; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 190 CTGGCGATTTCATGTGCGTTAAGGAGAAATTAGTGTCTCGGTAAACGGTTTCGACGAA 249

Db 527 CTGGCTCAAAGACTCTGGTTGGAGTCGAATACATTGTGAACTTACATCCCTTACACCTGTT 586

QY 250 CTGATATGACTCTGAATACATTAGTGAACCTTACATCCCTTACACCTGTT 309

Db 587 CTGATATGACTCTGAATACATTAGTGAACCTTACATCCCTTACACCTGTT 646

QY 310 CCCACCTTCCCATGTGAGATAAGTCGCAAAGTTCTCGGAAGCCCTCGTAAAGGT 369

Db 647 CAACTTCTCATGTATGTGTGTATGGTGTAAATGAGCAGTGTCAATTAAAGTT 706

QY 370 CAACTGGACGAAAACCTTGATATGAGTAAAGTGAATTAATAGAGAAAGAA 429

Db 707 CCAAGAAACCCCATTCAGCTGAATGTAGACCGCATGCTGAACCTGAAACTAGAA 766

QY 430 AAACCGTCTCGGTACTTTGCTACCCAAACACCCACGGAAACCTCTTCCAGG 489

Db 767 AACCCAAATGCCATTCACCTCTCCACATCCGAGTATCACGGAG 826

QY 490 GCAAAGATTGAGGAGATAAGAAACAGGGGTTTCTGTTAATAGACGAGCTACTAT 549

Db 827 GACGATCTGTCAGATTCTGAAATGCCAAATCTGTTCTAGATGAGCTACATC 886

QY 550 CATTACCGGAG 562

Db 887 GAATTCTAGGGAG 899

Search completed: April 15, 2003, 15:15:13
Job time : 3628 secs

Result No.	Score	Query Length	DB ID	Description
1	1065	100.0	1065	AAT78778
2	90.8	8.5	13732	AAA81484
3	90.8	8.5	349980	AAE21561
4	90.8	8.5	837096	AAA81489
c	5	60.7	1664376	AAV2109
c	6	46.2	4.3	1085
c	7	42.2	4.0	1092
c	8	37.6	3.5	588
c	9	37.6	3.5	588
10	37.6	3.5	588	AAT07876
11	37.6	3.5	588	AAT33743
12	37.6	3.5	588	AAI16517
13	37.6	3.5	588	AAI39468
14	37.6	3.5	588	AB508616
c	15	36.8	3.5	9835
c	16	36.4	3.4	6478
c	17	36.4	3.4	6478
c	18	36.4	3.4	6478
c	19	36.4	3.4	6626
c	20	36.2	3.4	6804
c	21	35.8	3.4	560
c	22	35.8	3.4	1098
c	23	35.8	3.4	1221
c	24	35.8	3.4	1221
c	25	35.8	3.4	10732
c	26	35.8	3.4	349980
c	27	35.4	3.3	4590
c	28	35.2	3.3	1104
c	29	35.2	3.3	1350
c	30	35	3.3	5621
c	31	35	3.3	24187
c	32	34.8	3.3	4213
c	33	34.8	3.3	4213
c	34	34.8	3.3	6631
c	35	34.8	3.3	6631
c	36	34.6	3.2	263
c	37	34.6	3.2	1341
c	38	34.6	3.2	24453
c	39	34.6	3.2	8559
c	40	34.6	3.2	19443
c	41	34.4	3.2	42
c	42	34.4	3.2	375
c	43	34.4	3.2	375
c	44	34.4	3.2	1225
c	45	34.4	3.2	12616
c	46	34.4	3.2	18
c	47	34.4	3.2	AT89256
c	48	34.4	3.2	AT89255
c	49	34.4	3.2	AT89254
c	50	34.4	3.2	AT89253
c	51	34.4	3.2	AT89252
c	52	34.4	3.2	AT89251
c	53	34.4	3.2	AT89250
c	54	34.4	3.2	AT89249
c	55	34.4	3.2	AT89248
c	56	34.4	3.2	AT89247
c	57	34.4	3.2	AT89246
c	58	34.4	3.2	AT89245
c	59	34.4	3.2	AT89244
c	60	34.4	3.2	AT89243
c	61	34.4	3.2	AT89242
c	62	34.4	3.2	AT89241
c	63	34.4	3.2	AT89240
c	64	34.4	3.2	AT89239
c	65	34.4	3.2	AT89238
c	66	34.4	3.2	AT89237
c	67	34.4	3.2	AT89236
c	68	34.4	3.2	AT89235
c	69	34.4	3.2	AT89234
c	70	34.4	3.2	AT89233
c	71	34.4	3.2	AT89232
c	72	34.4	3.2	AT89231
c	73	34.4	3.2	AT89230
c	74	34.4	3.2	AT89229
c	75	34.4	3.2	AT89228
c	76	34.4	3.2	AT89227
c	77	34.4	3.2	AT89226
c	78	34.4	3.2	AT89225
c	79	34.4	3.2	AT89224
c	80	34.4	3.2	AT89223
c	81	34.4	3.2	AT89222
c	82	34.4	3.2	AT89221
c	83	34.4	3.2	AT89220
c	84	34.4	3.2	AT89219
c	85	34.4	3.2	AT89218
c	86	34.4	3.2	AT89217
c	87	34.4	3.2	AT89216
c	88	34.4	3.2	AT89215
c	89	34.4	3.2	AT89214
c	90	34.4	3.2	AT89213
c	91	34.4	3.2	AT89212
c	92	34.4	3.2	AT89211
c	93	34.4	3.2	AT89210
c	94	34.4	3.2	AT89209
c	95	34.4	3.2	AT89208
c	96	34.4	3.2	AT89207
c	97	34.4	3.2	AT89206
c	98	34.4	3.2	AT89205
c	99	34.4	3.2	AT89204
c	100	34.4	3.2	AT89203
c	101	34.4	3.2	AT89202
c	102	34.4	3.2	AT89201
c	103	34.4	3.2	AT89200
c	104	34.4	3.2	AT89209
c	105	34.4	3.2	AT89208
c	106	34.4	3.2	AT89207
c	107	34.4	3.2	AT89206
c	108	34.4	3.2	AT89205
c	109	34.4	3.2	AT89204
c	110	34.4	3.2	AT89203
c	111	34.4	3.2	AT89202
c	112	34.4	3.2	AT89201
c	113	34.4	3.2	AT89200
c	114	34.4	3.2	AT89209
c	115	34.4	3.2	AT89208
c	116	34.4	3.2	AT89207
c	117	34.4	3.2	AT89206
c	118	34.4	3.2	AT89205
c	119	34.4	3.2	AT89204
c	120	34.4	3.2	AT89203
c	121	34.4	3.2	AT89202
c	122	34.4	3.2	AT89201
c	123	34.4	3.2	AT89200
c	124	34.4	3.2	AT89209
c	125	34.4	3.2	AT89208
c	126	34.4	3.2	AT89207
c	127	34.4	3.2	AT89206
c	128	34.4	3.2	AT89205
c	129	34.4	3.2	AT89204
c	130	34.4	3.2	AT89203
c	131	34.4	3.2	AT89202
c	132	34.4	3.2	AT89201
c	133	34.4	3.2	AT89200
c	134	34.4	3.2	AT89209
c	135	34.4	3.2	AT89208
c	136	34.4	3.2	AT89207
c	137	34.4	3.2	AT89206
c	138	34.4	3.2	AT89205
c	139	34.4	3.2	AT89204
c	140	34.4	3.2	AT89203
c	141	34.4	3.2	AT89202
c	142	34.4	3.2	AT89201
c	143	34.4	3.2	AT89200
c	144	34.4	3.2	AT89209
c	145	34.4	3.2	AT89208
c	146	34.4	3.2	AT89207
c	147	34.4	3.2	AT89206
c	148	34.4	3.2	AT89205
c	149	34.4	3.2	AT89204
c	150	34.4	3.2	AT89203
c	151	34.4	3.2	AT89202
c	152	34.4	3.2	AT89201
c	153	34.4	3.2	AT89200
c	154	34.4	3.2	AT89209
c	155	34.4	3.2	AT89208
c	156	34.4	3.2	AT89207
c	157	34.4	3.2	AT89206
c	158	34.4	3.2	AT89205
c	159	34.4	3.2	AT89204
c	160	34.4	3.2	AT89203
c	161	34.4	3.2	AT89202
c	162	34.4	3.2	AT89201
c	163	34.4	3.2	AT89200
c	164	34.4	3.2	AT89209
c	165	34.4	3.2	AT89208
c	166	34.4	3.2	AT89207
c	167	34.4	3.2	AT89206
c	168	34.4	3.2	AT89205
c	169	34.4	3.2	AT89204
c	170	34.4	3.2	AT89203
c	171	34.4	3.2	AT89202
c	172	34.4	3.2	AT89201
c	173	34.4	3.2	AT89200
c	174	34.4	3.2	AT89209
c	175	34.4	3.2	AT89208
c	176	34.4	3.2	AT89207
c	177	34.4	3.2	AT89206
c	178	34.4	3.2	AT89205
c	179	34.4	3.2	AT89204
c	180	34.4	3.2	AT89203
c	181	34.4	3.2	AT89202
c	182	34.4	3.2	AT89201
c	183	34.4	3.2	AT89200
c	184	34.4	3.2	AT89209
c	185	34.4	3.2	AT89208
c	186	34.4	3.2	AT89207
c	187	34.4	3.2	AT89206
c	188	34.4	3.2	AT89205
c	189	34.4	3.2	AT89204
c	190	34.4	3.2	AT89203
c	191	34.4	3.2	AT89202
c	192	34.4	3.2	AT89201
c	193	34.4	3.2	AT89200
c	194	34.4	3.2	AT89209
c	195	34.4	3.2	AT89208
c	196	34.4	3.2	AT89207
c	197	34.4	3.2	AT89206
c	198	34.4	3.2	AT89205
c	199	34.4	3.2	AT89204
c	200	34.4	3.2	AT89203
c	201	34.4	3.2	AT89202
c	202	34.4	3.2	AT89201
c	203	34.4	3.2	AT89200
c	204	34.4	3.2	AT89209
c	205	34.4	3.2	AT89208
c	206	34.4	3.2	AT89207
c	207	34.4	3.2	AT89206
c	208	34.4	3.2	AT89205
c	209	34.4	3.2	AT89204
c	210	34.4	3.2	AT89203
c	211	34.4	3.2	AT89202
c	212	34.4	3.2	AT89201
c	213	34.4	3.2	AT89200
c	214	34.4	3.2	AT89209
c	215	34.4	3.2	AT89208
c	216	34.4	3.2	AT89207
c	217	34.4	3.2	AT89206
c	218	34.4	3.2	AT89205
c	219	34.4	3.2	AT89204
c	220	34.4	3.2	AT89203
c	221	34.4	3.2	AT89202
c	222	34.4	3.2	AT89201
c	223	34.4	3.2	AT89200
c	224	34.4	3.2	AT89209
c	225	34.4	3.2	AT89208
c	226	34.4	3.2	AT89207
c	227	34.4	3.2	AT89206
c	228	34.4	3.2	AT89205
c	229	34.4	3.2	AT89204
c	230	34.4	3.2	AT89203
c	231	34.4	3.2	AT89202
c	232	34.4	3.2	AT89201
c	233	34.4	3.2	AT89200
c	234	34.4	3.2	AT89209
c	235	34.4	3.2	AT89208
c	236	34.4	3.2	AT89207
c	237	34.4	3.2	AT89206
c	238	34.4	3.2	AT89205
c	239	34.4	3.2	AT89204
c	240	34.4	3.2	AT89203
c	241	34.4	3.2	AT89202
c	242	34.4	3.2	AT89201
c	243	34.4	3.2	AT89200
c	244	34.4	3.2	AT89209
c	245	34.4	3.2	AT89208
c	246	34.4	3.2	AT89207
c	247	34.4	3.2	AT89206

PT used for producing enzymes, and hybridisation probes for a cDNA or genomic library
 PR XX
 PS XX
 Claim 1; Fig 7; 95pp; English.
 This genomic DNA encodes histidinol-phosphate aminotransferase VFS5/HPA (see AAW24254) of Aquifex VPS5, a strictly chemolithoautotrophic, marine eubacterium which grows optimally at 85°C and pH 6.8 in high salt medium. The VFS5/HPA sequence can be amplified from a pBuescript vector that contains the DNA by PCR (see AAT8794-95). The DNA can be used for recombinant probes of the aminotransferase and to develop hybridisation probes. Claimed thermostable transaminases and aminotransferases (AAW24248-57) can be produced from native or recombinant host cells for use with L- and/or D-amino acids for production of optically pure chiral compounds used in the pharmaceutical, agricultural and other industries. A method is claimed for transferring an amino group from an amino acid to an alpha-keto acid using a claimed enzyme.

Sequence 1065 BP; 344 A; 220 C; 252 G; 249 T; 0 other;

Query Match 100.0%; Score 1065; DB 18; Length 1065;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATACCCCAGAGGATTAAAGGAACTTGAAACCTTACAAAGCCGAGGTCACTCCGGCTCC 60

DB 1 ATGATACCCAGAGGATTAAAGGAACTTGAAACCTTACAAAGCCGAGGTCACTCCGGCTCC 60

QY 61 GTCAAGGCTTCCTCTAAGGAATTCAGACTTACGACTTCCGAGGATAAAACAAAGGGC 120

DB 61 GTCAAGGCTTCCTCTAAGGAATTCAGACTTACGACTTCCGAGGATAAAACAAAGGGC 120

QY 121 TAGAAGAATTAAAAGGTTCCCTGAAACAATAACCCAGAACCCGAAAGGGATA 180

DB 121 TAGAAGAATTAAAAGGTTCCCTGAAACAATAACCCAGAACCCGAAAGGGATA 180

QY 181 AAACGGGTTCTGGGATTTTTCGGGATTTTTCGGCTTAAGGAAGAAATTAGTCTCGTAACGGT 240

DB 181 AAACGGGTTCTGGGATTTTTCGGGATTTTTCGGCTTAAGGAAGAAATTAGTCTCGTAACGGT 240

QY 241 TCGGACGAACTTACATACCTACCTCACTAATAGCTATAGTTGAACTTTACACCGTTAAC 300

DB 241 TCGGACGAACTTACATACCTACCTCACTAATAGCTATAGTTGAACTTTACACCGTTAAC 300

QY 301 ATACCTGTTCCACCTTCCCATSTACAGGATAAATGCGGAACACTTCTGGAAACCCCTC 360

DB 301 ATACCTGTTCCACCTTCCCATSTACAGGATAAATGCGGAACACTTCTGGAAACCCCTC 360

QY 361 GTAANGTTCAACTGGACAAACTTGTATAGCTATAGCTAAAGAACTTGTAAATA 420

DB 361 GTAANGTTCAACTGGACAAACTTGTATAGCTATAGCTAAAGAACTTGTAAATA 420

QY 421 GAGAAAGAAAACCGGTTCTGGTACTTGTACCCAAACACCCACGGGAAACCTC 480

DB 421 GAGAAAGAAAACCGGTTCTGGTACTTGTACCCAAACACCCACGGGAAACCTC 480

QY 481 TTTCAGGGAAAGATGAGGAGATAAGAACAGGGGTTCCTGTTAGTACCAA 540

DB 481 TTTCAGGGAAAGATGAGGAGATAAGAACAGGGGTTCCTGTTAGTACCAA 540

QY 541 GCCTACTATCAATTACTCGGA GAAACCTTCTGGAAAGGGAGGATACG 600

DB 541 GCCTACTATCAATTACTCGGA GAAACCTTCTGGAAAGGGAGGATACG 600

QY 601 CTAGTTGAGGACACTTCAAAATCGGTTATGGCGACTTTAGGTAGGGATTATA 660

DB 601 CTAGTTGAGGACACTTCAAAATCGGTTATGGCGACTTTAGGTAGGGATTATA 660

QY 661 GGAAAGGGAAATCTCTAGAAATTAAAGGTGAGACTCCCTTAACGTGACCTCT 720

DB 661 GGAAAGGGAAATCTCTAGAAATTAAAGGTGAGACTCCCTTAACGTGACCTCT 720

PT - used for producing enzymes, and hybridisation probes for a cDNA or genomic library	Qy 721 CCCCTCTCAGGTGAAAGGTTCTCCACCGAGGAAGAACCTTAATGGAAAG 780
PR XX	Db 721 CCCCTCTCAGGTGAAAGGTTCTCCACCGAGGAAGAACCTTAATGGAAAG 780
PS XX	Qy 781 ATACAGGAGGTGTAACAGCGGAAAGGATGACGAACTTGAATAGAGGA 840
CC XX	Db 781 ATACAGGAGGTGTAACAGCGGAAAGGATGACGAACTTGAATAGAGGA 840
CC VFS5/HPA (see AAW24254) of Aquifex VPS5, a strictly chemolithoautotrophic, marine eubacterium which grows optimally at 85°C and pH 6.8 in high salt medium. The VFS5/HPA sequence can be amplified from a pBuescript vector that contains the DNA by PCR (see AAT8794-95). The DNA can be used for recombinant probes of the aminotransferase and to develop hybridisation probes. Claimed thermostable transaminases and aminotransferases (AAW24248-57) can be produced from native or recombinant host cells for use with L- and/or D-amino acids for pharmaceutical, agricultural and other industries. A method is claimed for transferring an amino group from an amino acid to an alpha-keto acid using a claimed enzyme.	Qy 841 GTTAGGTTTCCTCGGACTTACGTTCTGCTGTTAGAACGCTTACCCGCC 900
CC	Db 841 GTTAGGTTTCCTCGGACTTACGTTCTGCTGTTAGAACGCTTACCCGCC 900
CC	Qy 901 GAGTTTACAGGACTACTGAAAGGATGCTCTCGTAAAGGAAACCTATACATGAA 960
CC	Db 901 GAGTTTACAGGACTACTGAAAGGATGCTCTCGTAAAGGAAACCTATACATGAA 960
CC	Qy 961 GGACTCCAAAGTGCCTAGGGTAAGGTAAAGGAAAGAACACAAAGTTCTG 1020
CC	Db 961 GGACTCCAAAGTGCCTAGGGTAAGGTAAAGGAAAGAACACAAAGTTCTG 1020
XX	Qy 1021 GAAGCACTGGAGGAGTAAATCCCTTCAGCTCTTAA 1065
XX	Db 1021 GAAGCACTGGAGGAGTAAATCCCTTCAGCTCTTAA 1065
RESULT 2	
AAA81484 standard; DNA; 13732 BP.	
ID XX	
AAA81484;	
AC XX	
DT 04-DEC-2000 (first entry)	
XX N. meningitidis partial DNA sequence gnm_32 SEQ ID NO:32.	
DE ID AAA81484 standard; DNA; 13732 BP.	
XX XX	
Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic; KW	
antigen; vaccine; diagnosis; infection; antibacterial; identification; Menigococcus B; MenB; ds.	
XX OS	
Neisseria meningitidis.	
XX XX	
WO200022430-A2.	
XX XX	
PD 20-APR-2000.	
XX XX	
99WO-US23573.	
PF 08-OCT-1999;	
XX XX	
98US-0103794.	
PR 09-OCT-1998;	
PR 30-APR-1999;	
99US-0132068.	
XX XX	
(CHIR) CHTRON CORP.	
XX PR 08-OCT-1999;	
XX PR 09-OCT-1998;	
PI 09US-0132068.	
PI Masignani V, Hickey E, Peterson J, Tettelin H, Ventre JC;	
PI Ratti G, Galeotti C, Mora M, Scarlato V;	
PI Rappuoli R, Pizza M;	
XX DR 2000-318079/27.	
XX XX	
Isolated nucleotide sequences of Neisseria meningitidis which can be used in the diagnosis and treatment of N. meningitidis infection and other Neisseria infections, for example, N.gonorrhoea -	
XX XX	
Claim 7; Page 589-593; 1760pp; English.	
XX XX	
The present invention describes methods of obtaining immunogenic proteins from Neisseria genomic sequences. Neisseria meningitidis genomic DNA represent specifically claimed Neisseria meningitidis genomic DNA sequences; AAA81303 and AAB25663 represent Neisseria DNA sequences and their corresponding proteins; AAA81254 to AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the isolation of Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF	

CC sequences, which are all used in the exemplification of the present invention. The nucleic acid sequences, protein sequences, and antibodies against them, can be used in the manufacture of a composition. The composition can be used as a medicament (or in the manufacture of a medicament) for treating, preventing or diagnosing infection due to Neisserial bacteria. For example, some of the identified proteins could be components of vaccines against Meningococcus B; against all serotypes; and/or against all pathogenic Neisseriae. Identification of sequences from the bacterium will also facilitate production of biological probes, particularly organism-specific probes. Attempts to make efficacious Menigococcus B vaccines have failed mainly due to antigen tolerance. Multivalent vaccines have also been tried but none have successfully overcome antigenic variability. The provision of further, complete sequences may provide an opportunity to identify secreted or surface exposed proteins that may be presumed targets for the immune system and which are not antigenically variable or at least more conserved than other more variable regions.

XX Sequence 13732 BP; 3488 A; 3771 C; 3357 G; 3116 T; 0 other;
 XX Query Match 8.5%; Score 90.8%; DB 21; Length 13732;
 XX Best Local Similarity 47.7%; Pred. No. 8 9e-17;
 XX Matches 331; Conservative 0; Mismatches 357; Indels 6; Gaps 2;
 XX Db 12654 AAGCATTAGTCGCGATCATCGATCTCGACATTCGGACTCGCGAACACGT 12713
 XX Qy 182 AAGCGGTCTCGGGATTTTCGGCTTAAGGAAGAAAATTAGTGTCTCGTAACCGGT 241
 XX Db 12714 CGGAGGAACCTGATGATACTGATCTACGATGCTGACCCGCAACCGGGCGGAATGTGG 12773
 XX Qy 302 TACCTGTTOCCCCACCTTCCCATGAGAATAAGTGCAAAACTCTCGGAAGACCCCTCG 361
 XX Db 12774 CAGCGGAACCCAGTTTCGCNATGATCGGCCACAAACGCCGCTGTAGGGATGGATTAGT 12833
 XX Qy 362 TAAGGTTAACCTGACGAAACCTTGTATAGATAGAAGAAAGTAAATAG 421
 XX Db 12834 TCGGGTTCACGAAACGGAGATTTCACCTCAACCTGCCCGCTCTCGAAGGCCGTC 12893
 XX Qy 422 AGARAGAAAAACCCGTTCTGGGTACTCTGGTTACCCAAACACCCACGGAAACCTCT 481
 XX Db 12894 GGARACACGCCCTGCCCAGCTTACCCCAACACCCACCGCTATGCT 12953
 XX Qy 482 TTTCAGGGAAAGATGA--GGAGATAAGAACAGGGGTCTCTGTAAATAGAC 538
 XX Db 12954 TCAGCGGTGGCAAAUTCAAGCGGTCTACGAGCTTCAGACGGCTCGTGTGATG 13013
 XX Qy 539 AAGCCTACTATCATTACTCAGGAAACCTTCTGGAAAGACCGGTCAAAGGGAGATA 598
 XX Db 13014 AACGCTAACGGGAAACGCTTCAGGCAAGCTTCAGGCAAGGAGATTCCCAAAC 13073
 XX Qy 599 CGGTAGTTGAGGACACTTCAAATCGGTATGGGAGTTAAGGTAGGGATTAA 658
 XX Db 13074 TGATAGTTAGCACCTCAGCAAAATCGGTTTGGGACTATCGTTATGGG 13133
 XX Qy 659 TAGGAAAGGGAAATCTCGCAAAATTACAAGGTGAGGACTCCCTAACGTTAC 718
 XX Db 13134 CAGCTGCCGAAGTCATGCCGAACATGCCAAATCGGTTTGGGACTATGCC 13193
 XX Qy 719 ACCCTCTCAGGTGATGGCAAAAGTCTCCCTAACGGAGAGAAATTCTAATGAAA 778
 XX Db 13194 AATTGAGGCTGACCCTGCACACTCGCCCTGCGGACTAACGGCAATTCTCT 13250
 XX Qy 779 AGATACAGGGAGTTGTAACAGCGGAAAGATGATGACGCCAAATGAAAG 838
 XX Db 13251 ACATCGACAGCTGAAAGCAGCGGAAACCTGCCCTAACATGGGAAATATGCC 13310
 XX Qy 839 GAGTTGAGTTTCCGAGTAAGGTAAACTCTCT 872
 XX Db 13311 GTCTGAACACCTTCAAGTCAGGAAACTCAT 13344

RESULT 3
 ID AAF21611 standard; DNA; 349980 BP.
 XX
 XX
 XX
 AC AAF21611;
 DT 13-MAR-2001 (first entry)
 DE *Neisseria meningitidis* B nucleotide sequence SEQ ID NO:112.
 XX
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 KW ds.
 XX
 OS *Neisseria meningitidis*.
 XX
 PN WO200066791-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 08-MAR-2000; 2000WO-US050938.
 XX
 PR 30-APR-1999; 99US-0132068.
 PR 08-OCT-1999; 99WO-US23573.
 PR 28-FEB-2000; 2000GB-0004695.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Pizza M, Hickey E, Peterson J, Tettelin H, Ventter JC,
 PI Galeotti C, Mora M, Ratti G, Scarlato V, Rappoli R;
 PI Frazer CM, Grandi G;
 XX
 DR WPI; 2000-647603/62.
 XX
 PT *Neisseria meningitidis* B full length genome sequence and open reading frames are used to detect, treat and prevent *Neisseria* infections -
 XX
 PS Claim 7; Appendix A; 692PP; English.
 XX
 CC The present invention describes the full length genome of
 CC *Neisseria meningitidis* B (NMB). The sequences in AAF21544 and AAF21607
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the
 CC sequence was too long to go in a record on its own and end of each
 CC sequences which overlap each other at the beginning and end of each
 CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
 CC the beginning of AAF21607, the last 49980 bp of AAF21607, and so on). AAF21545 to
 CC AAF21588 encode the NMB proteins given in AAB58550 to AAB58593, and AAF21589 to
 CC AAF21606 represent PCR primers which are used in the exemplification of
 CC the present invention. The NMB genome and fragments from it have
 CC antibacterial activity, and can be used in vaccines and gene therapy.
 CC *Neisseria* nucleic acids, proteins and/or antibodies which binds to the
 CC proteins can be used in compositions for treating or preventing infection
 CC due to *Neisseria* bacteria or as a diagnostic reagent for detecting the
 CC presence of *Neisseria* bacteria or of antibodies raised to *Neisseria*
 CC bacteria. Computers, computer memory, computer storage medium or computer
 CC databases can be used in a search to identify open reading frames (ORFs)
 CC or coding sequences within the NMB genome. The DNA sequences provide
 CC further opportunities to find antigenic or immunogenic proteins which are
 CC more effective in vaccines than the outer membrane proteins currently
 CC used.
 XX
 SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 other;
 SQ Query Match 8.5%;
 SQ Best Local Similarity 47.7%;
 SQ Matches 331; Conservative 0; Mismatches 357; Indels 6; Gaps 2;

Qy	242	CGGACGAACTCATATACTACCTCTCAATAGCTTATAGCTGAACTTTAATACCCGTTTACA	301
Ddb	144679	CGGACGAACTGATACTAGTTCATCACGATGCTAACCGGGGAATGTGG	144738
Qy	302	TACCTGTCCCACCTTTCCTGATGAGATAAGTGGCAAAGAACCCCCCTCG	361
Ddb	144739	CAGGCCGAACCCAGTTCTCATGATCCTGAGCTAACCGGGCAAGACCTCTG	144798
Qy	362	TAAGGTTCAACTGGACCAAACCTTGATATAGACTTAGAAAGAAGTATTGAAATAATAG	421
Ddb	144799	TGGGGCTTCCACTGAACGGAGATTTCACCCCTAACCTCCGGCTCTCGAAGGCCCTCA	144858
Qy	422	AAGAAAGAAAAACCGGTTCTGGGTACTTGTCTACCCAAACAGGGAAACCTCT	481
Ddb	144859	GGAAAACCGGCCCTGGCCCTGACCTTATGCCTAACCCCAACACCCCAACCGGGTGTATCT	144918
Qy	482	TUTCCAGGGAAAGATTCA--GGAGATAAGAACACGGGTGTTTCTGTTATAGCG	538
Ddb	144919	TCAGCGTGGCGAAATCAGCGCTCATCGAACCTCTCGACGGCATCTCTGCTGTGTG	144978
Qy	539	AAGCCTACTATCATTCATTCTCCGGAAAACCTTCTGGAAAGACGGCTCAAAAGGAAGATA	598
Ddb	144979	AAGCCTACGGCGCATCAACGGGACACSTCTCCGGCAGGAGATCCCCAAC	145038
Qy	599	CGSTAGTTTGGACACTTCAAAACGGTTATGGCGAGTTAAGGGTAGGGATTAA	658
Ddb	145039	TGATAGCTTACACCCCTCAGGAAATGGTTTGGGACTCTGGTTATGGGG	145098
Qy	659	TAGGGAAAGGGAAATGCTCTCGAAATTAAACAGGTAGAGCTCCCTCAAGTGCCT	718
Ddb	145099	CAGGCTGGCCCGAGTCACTGGGAACITGCAAAATCTGCGCCCTAACATAACC	145158
Qy	719	ACCCCTCTAGGTGATGGCAAAGTTCTCAAGGGGAAGAGAAATTCCTAATGGAAA	778
Ddb	145159	AATGAGGCTGACCACTGCGCAAACTGCCCTGGGACTAGGGCATTATCTCT	-GCCA 145215
Qy	779	AGATAAGGGTGTGACAGGGAGAAGAGTACAGGGAAATGAGAAATAGTAAAG	838
Ddb	145216	ACATCGACAGCCGTAAGAACAGGGAATGGCAAAATAATGCC	145275
Qy	839	GAGTGGAGTTTCTCGAGTAAGGCTTAACCTCTT	872

RESULT 4
 AAA81489 standard; DNA; 837096 BP.
 ID AAA81489
 XX
 XX
 AC AAA81489;
 AC
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Neisseria meningitidis B; MenB; qs.
 XX
 Neisseria meningitidis.
 XX
 WO2000022430-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US23573.
 XX
 PR 09-OCT-1998; 98US-0103794.
 PR 30-APR-1999; 99US-0132068.
 XX
 PA (CHIRON CORP.
 PA
 PA
 PA
 PI Frazer CM, Hickey E,
 Peterson J, Tettelin H,
 Venter JC;

Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
IPI
Rappuoli R, Pizza M;
PCT
WPI; 2000-318079/27.

Isolated nucleotide sequences of *Neisseria meningitidis* which can be used in the diagnosis and treatment of *N. meningitidis* infection and other *Neisseria* infections, for example, *N.gonorrhoea* -

Claim 7: Page 629-865; 1760pp; English.

The present invention describes methods of obtaining immunogenic proteins from *Neisseria* genomic sequences. AAA81453 to AA82414 represent specifically claimed *Neisseria meningitidis* genomic DNA sequences; AA81260 to AAA81303 and ARB25660 to AAB25663 represent *Neisseria* DNA sequences and their corresponding proteins; AAA81554 to AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the isolation of *Neisseria meningitidis* DNA sequences; and AA813122 to AA81452 represent *Neisseria meningitidis* MenB polynucleotide OPR sequences, which are all used in the exemplification of the present invention. The nucleic acid sequences, protein sequences, and antibodies against them, can be used in the manufacture of a composition. The composition can be used as a medicament (or in the manufacture of a medicament) for treating, preventing or diagnosing infection due to *Neisseria* bacteria. For example, some of the identified proteins could be components of vaccines against *Meningococcus B*; against all serotypes and/or against all pathogenic *Neisseriae*. Identification of sequences from the bacterium will also facilitate production of biological probes, particularly organism-specific probes. Attempts to make efficacious *Meningococcus B* vaccines have failed mainly due to antigen tolerance. Multivalent vaccines have also been tried but none have successfully overcome antigenic variability. The provision of further, complete sequences may provide an opportunity to identify secreted or surface exposed proteins that may be presumed targets for the immune system and which are not antigenically variable or at least more conserved than other more variable regions.

Query	Match	Score	No. 8.5%;	Score	No. 9.7%;	Score	No. 7.4e-16;	Score	No. 3.7e-3;
	Best Local Similarity	Local	Mismatches	Local	Matches	Local	Matches	Local	Matches
Qy	182	AAGCGGTTCTGGCGATTTCGGGTAAAGGANGAAAATTAGTTCGGTAACGGTT	241						
Db	209330	AAGGATTACGTGTCGGTCACTCCGACTGGCGCATCCGGCAGTCGGCAACGGT	209449						
Qy	242	CGGACGAACTCATATACTACCTCNAATAGCTATAGGTGAACTTACATACCGGT	301						
Db	209450	CGGACGAACTGATCACGATGTCGGCACAAGGGCGGCAATGTGG	209509						
Qy	302	TACCGTTCACCTTCACCTTGATGAGATAAAGTGCAGAAAGTGTCTGGAGACCC	361						
Db	209510	CAGCGAACCCAGTTCTGTCATGTCAGGCCACACAGCGCGCTGTACGGCATGGTATG	209569						
Qy	362	TAAGGTTCAACTGGACAAAACCTTGTATACTAGCTAGAAAGAATGATGAATTA	421						
Db	209570	TGCGCGTTCACTTACCGGATTCACCTCAACCTCTAACCTGCGCCGTCCTCGA	209629						
Qy	422	AGAAGAAAACCCTGTTGGGTACTTGTCTACCCAAAACAACCCCACGGGAAACCTCT	481						
Db	209630	GGAAACACCGCCCTGGCCCTGACCCCTTATCGCTTACCCCAACCCCGCTGTAGT	209683						
Qy	482	TTCAGGGAAAAGATTGA --GGAGATAAGAAACAGGGTTTGTGTAAATGAGC	538						
Db	209690	TCAGCGTGTGCGAAATTGAGACGTCAGGGCATGTCGTGTGATG	209744						
Qy	539	AAGCTACTATCATTTACTCGGAGAAAACCTTCTGGAAAGACGGCCTAAAGGAAAGATA	598						
Db	209750	AAGCTACTGGCGCATCAACGGCGACAGTCCTGGCAGGGAGATTCCCAACC	209800						
Qy	599	CGTAGTTGAGGACATTCAAAATCGTATGGGAGTTAAGGTAGGGATTTAA	658						

Db	209810	TGATAGTCCTACGCCACCCTAGCAAATCGGTTTGGGACTCGTATGGTTATGCCG 209869	CC in AAV21211), can be used in the identification of <i>M. jannaschii</i> genome fragment.
Qy	659	TAGGGAGGGAAATCGTCAGAAAATTAAACAGGTGAGACTGCCCTCAACGTGACCT 718	CC CC Sequence 1664976 BP; 568133 A; 264649 C; 258701 G; 573392 T; 101 other; XX SQ
Db	209870	CAGGCTGCCGGAAGTCATGGCAGACTAACCTCGCCCTAACATAATGAAAC 209929	Query Match 5.7%; Score 60.2; DB 19; Length 1664976;
Qy	719	ACCCCTCTCAGGTGATGCCAAAAGTCTCTCAGGGAGGAAGAATTCTATGGAA 778	Best Local Similarity 47.7%; Pred. No. 2.9e-06;
Db	209930	AATTGAGCCTGACCCTAACCTCGCCCTGCGCAGTACAGGCTATTCTCTC - - GACA 209986	Matches 479; Conservative 0; Mismatches 493; Indels 33; Gaps 9;
Qy	779	AGATACAGGGGTTGTAACAGGGAGAAGGATGACGGAAATAGAGAATAGAGG 838	Qy 62 TAGGCTTCCTCTAAGCAGATTCCCTACTCGACTTTCCGAGGATAAAACAAAGGGCCT 121
Db	209987	ACATCGACACCTGAAACAGCGAACGGATGTTCCCGAATGGCAAAATAATGCC 210046	Db 586070 TAAACATGGTCTCATGAAAATCTGGGACCACCTCCAAAATAAAAGAAAATTT 586011
Qy	839	GAGTGAGGTTTCGAGTAAGGCTAACTTCAT 872	Qy 122 TAGAGAAATTAAAAGGTTCCCTTGACAAATAACCCAGACCCGAGCGAAAGCTAA 181
Db	210047	GTCGAACACCTTCAAGTCAGGCAAACATTCAAT 210080	Db 586010 TAGATGAAATTGACAAATTC - - ACCAAATATCAGAGCCAGTAATCCAAATTAA 585957
RESULT 5			Qy 182 AACCGGTTCTCGGATTTCGGGTTAAGGAAATAATTTGAACTTGTGAGGAGATG 238
ID	AAV21209/c		Db 585956 TGAAAGAGTTAACAAATTGAACTTGTGAGGAAACATAATTGTTGAGGAGATG 585897
XX	AAV21209 standard; DNA; 1664976 BP.		Qy 239 GTTCGGAGAACTCATATACTACCTCTCAATAGCTTATAGGTGAACCTTACACCGTT 298
AC	AAV21209;		Db 585896 GAGCTGTGAGATTAGACAAATTTAGAACGTTGTGTGATGAGGATGAGGTTA 585837
XX			Qy 299 ACATACGTGTCACCTTCCCATTAGCTGATAAGTGTGAAAGTCTGGAAACCCCC 358
DT	10-NOV-1998 (first entry)		Db 585836 TAATTCAAATTCCAAATTAGAGTTACGAAACAACTCCAAATGCCTAAA 585777
XX			Qy 359 TCGTAAGGTTCAACTGGACCAAACCTTGTATAGCTAGTAAAGAATGTTGATGTTAA 418
DE	Methanococcus jannaschii circular chromosome.		Db 585776 TAAATATGCTAAATTGATAAGGAAAGACTTAAATGTTGAAAGTGTGTCCTAA 585717
XX			Qy 419 TAGA -- -GAAAGAAANACCCCTTCTGGGACTCTGTTACCCAAACAAACCCAGGGAA 475
KW	Methanococcus jannaschii; methanogenic archaeon; circular chromosome; genome; autotrophic; extrachromosomal element; identification; ds.		Db 585716 ATAATATAACAGATAAAACGAAGGTTATTTCCTGACTCCAAATACTCCACAGGAA 585657
KW			Qy 476 ACCTCTTTCCAGGGGAAAGTTGA -- GGAGATAAGAAACAGGGTTTCTGTTGTA 532
OS	Methanococcus jannaschii.		Db 585656 ATATATAGAAAATAGAGAATGAGAGGTTATCATGAAACAGCGTTTACTGGTTA 585597
XX			Qy 533 TAGACCAAGCCCTACTATCATTACPCCGGAAACCTTT --- CTGGAAAGACGCCCTCA 586
PN	WO907830-A2.		Db 585596 TTGACCATGACATGACATGAGTGTGACTCTAAAACCCCTG 585537
XX			Qy 587 AAAGGAAAGATACGGTACTGTTGAGGACACTTCAAATTC -- GGATGGGAGSTTAA 643
PD	26-PEB-1998.		Db 585536 AATATCATAATGTTATGTTTAAGAACCTTCTAAAGGTTTACGAGGAAATGA 585477
XX			Qy 644 GGTTAGGGATTTAATGGAAAGGGGAAATTAACAGGTGACTCTCAAATGAGATG 703
PF	22-AUG-1997; 97WO-US14900.		Db 585476 CGTCACGTGACCTTCCCTCTAGGTGATGGTTGCAAATTAAGCTTCACGGAAAG 585417
XX			Qy 704 CGTCACGTGACCTTCCCTCTAGGTGATGGTTGCAAATTAAGCTTCACGGAAAG 763
PR	22-AUG-1996; 96US-0024428.		Db 585416 TATITAGCTTAACANGTTAAGTCAGTAACTGATTAAAGAT -- AGAG 585360
XX			Qy 764 AATTCTTAATGGAAAGATAACGGAGGTGTAAGGAGCTTCAAGGAAAG 823
PA	(GENO-) INST GENOMIC RES.		Db 585359 AATTCTTGTAAAGATGTGTTAGAGGAAATTAAAAGTAGAGATGCTCTAACATGGAT 585300
(UNI.)	UNIV ILLINOIS FOUND.		Qy 824 TGARGAAAATAGAGGAGTGTGAGTTTCCGAACTAAGCTAACT - - TCTTGCTTCTCA 880
(UJO)	UNIV JOHNS HOPKINS SCHOOL MEDICINE.		Db 585299 TGAAGAGTTAACGATTAATAAAGTTATCATGAGAATGAGCTTATGTTGTTA 585240
XX			Qy 881 GAACCCCTAACCGCCACAGGTTATCAGGAGCTACTGAAAGGGGATGTCCTGCTCA 940
Bult CJ, Smith HO, Venter JC, White OR, Woese CR;			Db 585239 TAAAAACAAATGAAAGCAAAGAAATTTGTGAGGAACTTAAAGGTATGTTA 585180
XX			Qy 941 GGAAGGGTACCTTATGGAGG -- ACTCCAAAGTGCCTCAGGGTAAAGGGAAAC 997
DR	WPI, 1998-169145/15.		Db 585179 GAGATTGCACATCOTTGATGGTTAGGGATAATATGTTAGAGTCAATGGGACGT 585120
XX			
PT	Complete genome sequence of methano-genic archaeon, <i>M. jannaschii</i> genome fragment.		
PT	jannaschii - useful in identification of <i>M. jannaschii</i> genome		
XX	Claim 13; Page 152-585; 614pp; English.		
CC	The present sequence represents the complete 1,66-megabase pair genome sequence of the Methanococcus jannaschii circular chromosome. The present invention describes <i>M. jannaschii</i> open reading frames from the genome sequence. The invention also describes a computer based system for identifying fragments of the <i>M. jannaschii</i> genome that are homologous to target nucleotide sequences, comprising: (a) data storage means comprising the nucleotide sequence of the 1664976, 58407 or 16550 bp sequence (see AAV21209, AAV2120 and AAV21211), or a nucleotide sequence at least 99.9% identical to it; (b) search means for comparing a target sequence to the nucleotide sequence of the data storage means to identify a homologous sequence; and (c) retrieval means for obtaining the homologous sequence. The method, which is based on whole genome random sequencing of an autotrophic archaeon <i>M. jannaschii</i> , the genome of which consists of 3 physically distinct elements, a large circular chromosome (the 1664976 bp sequence given in AAV21209), a large circular extra-chromosomal element (the 58407 bp sequence given in AAV2120), and a small circular extra-chromosomal element (the 16550 bp sequence given		

Qy 998 CGGAAGAAAACAAGTTCTGGAGCACTGGAGAGATAAA 1042
Db 585119 TGAAGGTTGAAAGTTTAAGATTTGAGGATTTAA 585075

RESULT 6

ABR73007 standard; DNA; 1085 BP.

XX ABR73007;
XX DT 13-AUG-2002 (first entry)
XX Bacillus licheniformis genomic sequence tag (GST) #298.
DE Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX Bacillus licheniformis.
OS WO200229113-A2.
XX PD 11-APR-2002.
XX PF 05-OCT-2001; 2001WO-US31437.
XX PR 06-OCT-2000; 2000US-0880598.
XX PR 27-MAR-2001; 2001US-279526P.
XX PA (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.
XX PI Berka R, Clausen IG;
DR WPI; 2002-416684/44.
XX PT Monitoring differential expression of several genes in first *Bacillus* cell relative to expression of same genes in one or more second *Bacillus* cells, by using substrate containing *Bacillus* genomic sequenced tag array -
XX PS Claim 4; SEQ ID NO 298; 200pp; English.
CC The invention describes a method of monitoring differential expression of genes in a first *Bacillus* cell relative to expression of the genes in other *Bacillus* cells, comprising hybridising labelled nucleic acid probes isolated from *Bacillus* cells to a substrate containing array of *Bacillus* genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first *Bacillus* cell relative to expression of the same genes in one or more second *Bacillus* cells. The method is useful for monitoring global expression of several genes from a *Bacillus* cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which *Bacillus* cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.

XX Sequence 1085 BP; 302 A; 264 C; 272 G; 247 T; 0 other;

Query Match 4.38; Score 46.2; DB 24; Length 1085;
Best Local Similarity 47.4%; Pred. No. 0.0014;
Matches 138; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

Qy 154 TACCCAGACCCSAAGGAAAGTAAAGGGTTCTGGGATTTCGGCGTTAAC 213
Db 181 TATCTGAGGGTATAGTGGSCTTGGAAAGTTGCCTCCCATCTGGCTAAAT 240

Qy 214 GAAGAAATTAGTACTCGTAACTGGTGGAACTCATACCTCTCAATAGCT 273
Db 241 GAACGARATCATCCTGGAAACGGAGCTCATCAAAATTCCCGTTCT 300

Qy 274 ATAGGTGAACCTATACCCGTTACATACCTGTTCCACCTTCCATGACGATA 333
Db 301 TTGGGGACCCGCTCAAAACACTGTAATGGGAATGGACATTTCACAAATCACGCCAC 360

Qy 334 AGTGGAAAGTTCTGGAAAGACCCCTGTAAGGTCAACTGACGAAACTTGTATA 393
Db 361 AACCTGTATCGAGGAGGGAGTCAAGAGAAAGAAAAACCGGTTCTCGGG 444
Qy 394 GACTGAAAGAAGTATGAAATATGAAAGAAAAACCGGTTCTCGGG 444
Db 421 GACCTGTATGCAATGCTGAGGGATGGTACAGACGAAAGTTGTCTGG 471

RESULT 7

AAT78780 standard; DNA; 1092 BP.

XX ID AAT78780;
XX AC AAT78780;
XX DT 18-FEB-1998 (first entry)
XX DE Ammonifex histidinol-phosphate aminotransferase KC4/HPA DNA.
XX Histidinol-phosphate aminotransferase; KC4/HPA; chiral compound;
KW KW ss.
XX OS Ammonifex degensii strain KC4.
XX PN WO979187-A1.
XX PD 14-AUG-1997.
XX PR 21-JAN-1997; 97WO-US01094.
XX PR 08-MAY-1996; 96US-0646590.
XX PR 09-FEB-1996; 96US-0599171.
XX DR WPI; 1997-415343/38.
XX PT New transaminase(s) and aminotransferase(s) derived from host cells used for producing enzymes, and hybridisation probes for a cDNA or genomic library
XX PT genomic library
XX PS Claim 1; Fig 9; 95pp; English.
XX This genomic DNA encodes histidinol-phosphate aminotransferase KC4/HPA (see AAW24256) of *Ammonifex degensii* KC4, a new chemolithoautotrophic, eubacterium which grows optimally at 70 deg C and pH 7.0 in low salt medium. The KC4/HPA sequence can be amplified from a pBluescript vector that contains the DNA by PCR (see AAT78798-99). The DNA can be used for recombinant production of the aminotransferase and to develop hybridisation probes. Claimed the thermostable transaminases and aminotransferases (AAW24248-57) can be produced from native or recombinant host cells for use with L- and/or D-amino acids for production of optically pure chiral compounds used in the pharmaceutical, agricultural and other industries. A method is claimed for transferring an amino group from an amino acid to an alpha-keto acid using a claimed enzyme.

Sequence 1092 BP; 217 A; 312 C; 357 G; 205 T; 1 other;	CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences .
Query Match 4.0%; Score 42.2; DB 18; Length 1092;	CC	i
Best Local Similarity 48.2%; Pred. No. 0.024;	CC	
Matches 119; Conservative 0; Mismatches 128; Indels 0; Gaps 0;	XX	
138 GGTTCCCTTGACAAATACCCAGACCAGAACGCAAAGAGTAAAGCGGTCTTGGGA 197	Sequence 588 BP; 274 A; 71 C; 111 G; 132 T; 0 other;	
168 GGAACCATGGACCTTAACCGAAGCTAACCGAAGCTAACGCAAGCCTGGTAA 227	Query Match 3.5%; Score 37.6; DB 22; Length 588;	
198 TTTCCTGGCCTTAAGGAAATTAGTCTCGTAAGGTTGGGAACUCATA 257	Best Local Similarity 50.6%; Pred. No. 0.45;	
228 GAAACCTGGAGTAGACCCGGCACGATCATCGTGGTTGGCTAACGGAACTACCA 287	Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;	
258 CTACCCCTCATATAGCTATAGGTGAACTTACATACCGTTACACCCGTTCCACCTT 317	QY 680 CAGAAATTAAACRAGGTGAGACTCCCTTCAACGTCACCTACCCCTCTGATGGCAA 739	
288 GATGCCCTCTTGGGOCCTGGGCCCCCGGCAGGGTTGATCCCTGTCCTACCT 347	Db 309 CAGGCCCTTAGCATCTGGGACAAATGTCACATTGTCAGTGTGATGTTA 368	
318 TCCCCATGTACCGAGATAAGTGGAAAGTCTCGGAAAGACCCCTGCTAAAGGTCAACTGGA 377	QY 740 AAGTTCTCTACGGGGAGAGAATTTCTTAATGAAAGATAAGGGGTGTAACAG 799	
348 TCCCCGCTATAGGCCCTGGACGGCTAACATCCAGATTTAAACATCCAAAGTAAAGTCCAAAGTA 407	Db 369 CAGAAGAAAGATGTTGGGGAGAAGGATTTGAAACAAATAAAGGCCAANACT 428	
378 CGAAAC 384	QY 800 ACCGAGAAAGGATGTCGACCAAATGAAAGAAAATAGAAGGAGGTGAGGTTTTCGGAGTA 859	
	Db 429 TCCCCAAATCTGTAAGGACATAAAATACATCCAAAGATTTAAGGAAGTCCAAAGTA 488	

Note: The sequence data for this patent did not form part of the printed specification, but was obtained from WIPO at <http://www.wipo.int/pctdb/pct.html>.

RESULT 10
AAK0787
AAK3.37.3;
XX DT XX
06-NOV-2001 (first entry)

Human bone marrow expressed single exon probe SEQ ID NO: 3400.
 DE Human; bone marrow; expressed exon; gene expression analysis; probe;
 XX
 KW Human; bone marrow; expressed exon; gene expression analysis; probe;
 mIcroarray; cancer; leukemia; lymphoma; myeloma; ss.
 KW

D1
 U3-NOV-2001 (first cut) Human brain expressed single exon probe SEQ ID NO: 7867.
 XX
 DE
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW encephalitis; cancer; ss.
 XX
 WO200157276 A2.
 XX
 PN
 XX
 PD
 PR
 XX
 Homo sapiens.
 XX
 XX
 XX
 XX
 09-AUG-2001.
 4

PR 26 MAY -1000; 2000US-0507456.
PR 30 MAY -1000; 2000US-0508408.
XX WO2001157275-A2.

30-JAN-2001 2001IWO-US00657.

	PR	PR	SG	SG	DR	DR
30-JUN-2000	2000S	0608	0608	0608	XX	XX
03-AUG-2000	2000S	0618	0618	0618	PI	Penn SG,
		0623	0623	0623		Hanzel DR,
		0626	0626	0626		Chen W,
		0628	0628	0628		Rank DR;

WPI : 2001-488900/53 .
XX DR
XX

Human genome-derived single exon nuclear probes useful for analyzing gene expression in human bone marrow PT PT

Example 4 : SEQ ID NO: 8300: 658pp + Sequence Listing; English.
PS
XX

The present study was designed to analyze gene expression in human marrow using cDNA probes derived from genomic sequences expressed in the human genome. The probes have been used to measure gene expression in bone marrow samples obtained from patients with various diseases.

Sequence Listing: English.
CC samples which may enable the improved diagnosis of
CC such as lymphoma, leukaemia and myeloma. The present
CC sequence is one of
CC

The present invention provides a number of single-exon nucleic acid sequences. The following table summarizes the characteristics of these sequences:

disorders such as Alzheimer's disease, epilepsy, and cancers. The present sequence is one of the probes of the mismatch repair system; it has been used to study the mismatch repair system in various organisms, including *S. cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, and *Escherichia coli*. The sequence is also used to study the mismatch repair system in *Escherichia coli*.

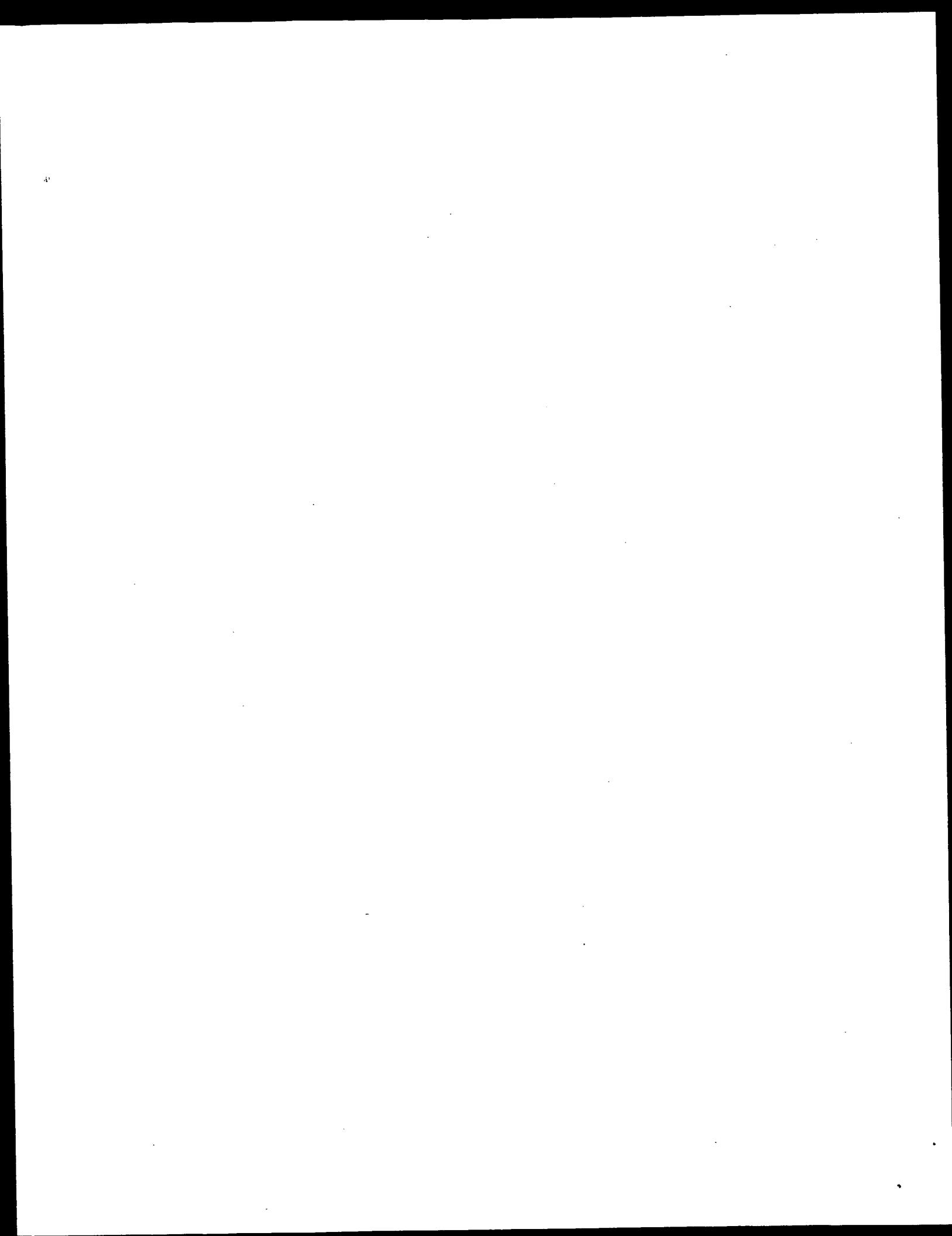
Sequence 588 BP; 274 A; 71 C; 111 G; 132 T; 0 other;

CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB5737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 9835 BP; 2538 A; 2196 C; 2372 G; 2729 T; 0 other;

	Query	Match	Score	Length	DB	Best	Local	Similarity	Pred.	No.	Mismatches	Indels	Gaps
Qy	53	CGGCCCTCGCTAGGGCTTCCCTACGAATCCCTACACTTCCCAAGGATAAAAC	3.5%	9835									
Db	3019	CGGCCCTACCAAGGCATTACCAACGAACTCACAATAATTAGCTAAAGAATAGAC	50.0%										
Qy	113	AAAGGCCCTAGAAAGATAAAAAAGGTCCCTGAAACAAATACCCAGACCCGAAGCA	36.8										
Db	2959	CAACGAAATAGCATTACGCACTTAAGGTTCTCCCGTATAACTAAAGAACATCAA	3.4										
Qy	173	AAGAGTTAAAGGGTTCTTGCGGATTTTTCGGCTTAAGGAGAAATTAGTCCTG	111										
Db	2899	TTTAGTAGAAAATGAAATATTATAGTATATATTATTATTATTATTATTTATA	111										
Qy	233	GTAA	236										
Db	2839	GTAA	2836										

Search completed: April 15, 2003, 14:15:26
 Job time : 1558 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 14:00:46 ; Search time 63 Seconds
(without alignments)
5184.298 Million cell updates/sec

Title: US-09-905-173-23
Perfect score: 1065
Sequence: 1 atgataccaggattaa.....cccttcaagtcctcttaa 1065

Scoring table: IDENTITY NUC
Gapext 1.0
Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cn2_6/podata/1/ina/5A_COMB.seq;*
2: /cn2_6/podata/1/ina/5B_COMB.seq;*
3: /cn2_6/podata/1/ina/6A_COMB.seq;*
4: /cn2_6/podata/1/ina/6B_COMB.seq;*
5: /cn2_6/podata/1/ina/PCRTS_COMB.seq;*
6: /cg2_6/podata/1/ina/backfiles1.seq;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1065	100.0	1065	1	US-09-599-171A-23
2	1065	100.0	1065	2	US-09-590B-23
3	1065	100.0	1065	3	US-09-069-226-23
4	1065	100.0	1065	4	US-09-012-184-23
c 5	61	5.7	7218	1	US-08-032-163-14
c 6	42.2	4.0	1092	2	US-08-046-190B-15
7	42.2	4.0	1092	4	US-09-412-184-35
8	34.4	3.2	12225	2	US-08-082-045-11
9	34.4	3.2	12225	4	US-09-396-040-11
10	34.4	3.2	12616	2	US-08-082-045-9
11	34.4	3.2	12616	4	US-09-396-040-9
c 12	33.2	3.1	1100	1	US-08-234-939-4
c 13	33.2	3.1	1100	1	US-08-558-065-7
c 14	33.2	3.1	1268	1	US-08-234-939-6
c 15	33.2	3.1	1268	1	US-08-558-065-6
c 16	33.2	3.1	1385	1	US-08-234-939-5
c 17	33.2	3.1	1385	1	US-08-558-065-5
c 18	33.2	3.1	1859	1	US-08-234-939-4
c 19	33.2	3.1	1859	1	US-08-558-065-4
c 20	33.2	3.1	1933	1	US-08-234-939-3
c 21	33.2	3.1	1933	1	US-08-558-065-3
c 22	33.2	3.1	3426	1	US-08-234-939-1
c 23	33.2	3.1	3426	1	US-08-558-065-1
c 24	33.2	3.1	3426	3	US-08-654-025-6
c 25	32.8	3.1	180	4	US-08-894-818B-31
c 26	32.8	3.0	1770	4	US-08-943-731-146
c 27	32.4	3.0	24183	4	US-08-943-731-3

ALIGNMENTS

RESULT 1
US-09-599-171A-23
; Sequence 23, Application US/08599-171A
; Patent No. 5814473
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CECCHI, BYRNE, BAIN, GILFILLAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599-171A
; FILING DATE: Concurrently
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28, 019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1065 NUCLEOTIDES
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: GENOMIC DNA
; US-08-599-171A-23
; Query Match 100.0%; Score 1065; DB 1; Length 1065;
; Best Local Similarity 100.0%; Pred. No. 0;
; Matches 1065; Conservative 0; Mismatches 0;
; Indels 0; Gaps 0;
; Qy 1 ATGATACCCAGAGGATAAGGAACCTGAAGGACTTACAGAGCTACTCCGGCTCC 60

OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/069, 226
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/599, 171
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: HERRON, CHARLES J.
 REGISTRATION NUMBER: 28, 019
 REFERENCE/DOCKET NUMBER: 331400-38
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INSTRUCTION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1065 NUCLEOTIDES
 TYPE: NUCLEIC ACID
 STRANDEDNESS: SINGLE
 TOPOLOGY: LINEAR
 MOLECULE TYPE: GENOMIC DNA
 US-09-069-226-23

Query Match Score 100.0%;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 ATGATACCCAGAGGATTAAAGGAACCTGTTACAGGGTCAACTGGCTTACCAACACCCACGGAAACCTC 480
Db	1 GAGAAAGAAAACCGTTCAGCTTGCTTAACCCACGGAAACCTC 480
Qy	1 TTTTCCAGGGAAAGATGAGGATATAAGAACAGGGTGTTCAGCTTACCCACGGAAACCTC 540
Db	1 GAGAAAGAAAACCGTTCAGCTTGCTTAACCCACGGAAACCTC 540
Qy	1 GCCTACTATCATACTCGGAAACCTTCTGGAGACGCCCTAAAGGGAGATAAG 600
Db	1 GCCTACTATCATACTCGGAAACCTTCTGGAGACGCCCTAAAGGGAGATAAG 600
Qy	1 GTAGTTTGGGAAAGATGAGGATATAAGAACAGGGTATGGTAGGTTAATAA 660
Db	1 GTAGTTTGGGAAAGATGAGGATATAAGAACAGGGTATGGTAGGTTAATAA 660
Qy	1 GGGAAAGGGAAATCTCAGRAATTAAACAGGTGAACTCCCCCTCAAGGTGACCTAC 720
Db	1 GGGAAAGGGAAATCTCAGRAATTAAACAGGTGAACTCCCCCTCAAGGTGACCTAC 720
Qy	1 CCCCTCTCAGGTGATGGCAAAGTTCTCACGGGGAGAGAAATTCTATGGAAAG 780
Db	1 CCCCTCTCAGGTGATGGCAAAGTTCTCACGGGGAGAGAAATTCTATGGAAAG 780
Qy	1 ATACAGGGAGTGTAAAGGGAGAAAGGAGTACAGGTGAACTCCCCCTCAAGGTGACCTAC 840
Db	1 ATACAGGGAGTGTAAAGGGAGAAAGGAGTACAGGTGAACTCCCCCTCAAGGTGACCTAC 840
Qy	1 GTTGAGGTTTCGAGTAAGGCTAAACTCTTGCTTTCAAGACGCCCTAACCCGCCAC 900
Db	1 GTTGAGGTTTCGAGTAAGGCTAAACTCTTGCTTTCAAGACGCCCTAACCCGCCAC 900
Qy	1 GAGTTTATCAGGACTACTGAAAGGATGTCGTCAGGAACGTATCTTACATGGAA 960
Db	1 GAGTTTATCAGGACTACTGAAAGGATGTCGTCAGGAACGTATCTTACATGGAA 960
Qy	1 GGACTCCAAAGTGCCTCAGGTTAACCGTAGGAAACCAAGTTCG 1020
Db	1 GGACTCCAAAGTGCCTCAGGTTAACCGTAGGAAACCAAGTTCG 1020
Qy	1 1021 GAAGCACTGAGGAGACTAAATCCCTCAAGCTCTTTAA 1065
Db	1 1021 GAAGCACTGAGGAGACTAAATCCCTCAAGCTCTTTAA 1065

RESULT 3
 US-09-069-226-23
 Sequence 23, Application US/09/069/226
 Patent No. 6013509
 GENERAL INFORMATION:
 APPLICANT: WARREN, Patrick V.
 TITLE OF INVENTION: TRANSAMINASES AND AMINO TRANSFERASES
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2

RESULT 5
US-08-232-463-14/c
Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: PALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300 .6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENET, Stephen A.
; REGISTRATION NUMBER: 29-768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZqpt-F1s
US-08-232-463-14

Query Match Similarity 5.7%; Score 61; DB 1; Length 7218;
Matches 7; Conservative 229; Mismatches 139; Indels 0; Gaps 0;

Qy 901 GAGGTTTACAGGAGCTACTGTAAAGGGATGTCCAGAACGTATCTACATGGA 960
Db 901 GAGGTTTACAGGAGCTACTGTAAAGGGATGTCCAGAACGTATCTACATGGA 960
Qy 961 GGACTCCAAGAGTCCTCAGGTAAGCTAGGAAACCGGAAGAAAACACAGTTCTG 1020
Db 961 GGACTCCAAGAGTCCTCAGGTAAGCTAGGAAACCGGAAGAAAACACAGTTCTG 1020
Qy 1021 GAAGGACTGGAGGAGATATAATCCCTTCAAGCTCTCTTAA 1065
Db 1021 GAAGGACTGGAGGAGATATAATCCCTTCAAGCTCTCTTAA 1065
Qy 593 AGATACGGTAGTTGAGGAACACTTCAAAATCCGTATGGCAGTTAAGGGTAGGGA 652
Db 1328 RRR 1269
Qy 653 TTTTAATAGGGAGGGAAATCGTCTCAGAAATTACAAGTGACACTCCCTCAACG 712
Db 1268 RRR 1209
Qy 713 TGACCTACCCCTCTCAGGTGATGGCAAAGGTCTCCCTCACGGGAAAGAGAAATTCTAA 772
Db 1208 RRR 1149
Qy 773 TGGAAAGATAACAGGGAGTGAACAGGCAGAAAGGAATGAGAAAAA 832
Db 1148 RRR 1089

RESULT 6
US-08-646-590B-35
; Sequence 35, Application US/08646590B
; Patent No. 592283
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,590B
; FILING DATE: 08-May-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; APPLICATION NUMBER: US/08/646,590B
; FILING DATE: 21-January-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hale, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 619/678-5099
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE: Coding Sequence
; NAME/KEY: Coding Sequence
; LOCATION: 1...1089
US-08-646-590B-35

Query Match 4 .0%; Score 42.2; DB 2; Length 1092;
 Best Local Similarity 48.2%;
 Matches 119; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Query Match 4 .0%; Score 42.2; DB 4; Length 1092;
 Best Local Similarity 48.2%;
 Matches 119; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Query Match 4 .0%; Score 42.2; DB 4; Length 1092;
 Best Local Similarity 48.2%;
 Matches 119; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 138 GTTCCCTGAAACAAATACCGAACGGGAAGAGTTAAAGCGGTCTTGGGA 197
 Db 168 GGACCACTGSCACCTTACCCAGAAAGGCTCAAGCTTACGCTGGCAGGGTAA 227

Qy 198 TTTCCTGGGAAATTACTTCGGTAACGGTACGGAGTCATAATACTGACATA 257
 Db 228 GAACTGGGATAGACCCGGACAGCATCATGGGTGCGGCTCAAGCGAAGTCACA 287

Qy 258 CTACCTCTCAAATAGCTATAGTGTAACTCCGTTAACATCCGTTACGTTAC 317
 Db 288 GATGCTCTTGCGCTGCGCTGTTGGCTGGGGGAGCTGGTACCTGGTACCT 347

Qy 318 TCCCATGTACGATAAGTTCGAAAGTCTCGGAAGACCCCTCTAAAGGTTAACCTGGA 377
 Db 348 TCCCCCTTGTGACGGCTCATGGGCTAATCCGTAAGTTCCTGAA 4.07

Qy 378 CGAAAC 384
 Db 408 GGAATC 414

RESULT 7
 US-09-412-184-35
 Sequence 35, Application US/09412184
 Patent No. 6268188

GENERAL INFORMATION:
 APPLICANT: Warren, Patrick V.
 TITLE OF INVENTION: TRANSMINASES AND AMINOTRANSFERASES
 NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson, P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: US
 ZIP: 92037

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows 95
 SOFTWARE: FASSEQ FOR Windows Version 2.0

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/412,184
 FILING DATE:

PRIOR APPLICATION DATA:
 CLASSIFICATION:
 APPLICATION NUMBER: PCT/US97/01094
 FILING DATE: 21-January-1997
 CLASSIFICATION:
 FILING DATE: 08-May-1996
 APPLICATION NUMBER: 08/599,171
 FILING DATE: 09-FEB-1996
 PRIOR APPLICATION DATA:
 REFERENCE/DOCKET NUMBER: 09010/017001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619/678-5070
 TELEFAX: 619/678-5059
 INFORMATION FOR SEQ ID NO: 35:
 SEQUENCE CHARACTERISTICS:
 REGISTRATION NUMBER: 38,347
 LENGTH: 1092 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA

FEATURE: Coding Sequence
 NAME/KEY: Coding Sequence
 LOCATION: 1...1099
 US-09-412-184-35

RESULT 8
 US-08-822-445-11
 Sequence 11, Application US/08822445
 Patent No. 595223
 GENERAL INFORMATION:
 APPLICANT: Kaplan, Jerry
 ADDRESS: Peru, Charles
 APPLICANT: Moore, Karen
 APPLICANT: More, Karen
 TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
 TITLE OF INVENTION: AND TREATMENT OF CHEDIAK HIGASHI SYNDROME
 NUMBER OF SEQUENCES: 32
 CORESPONDENCE ADDRESS:
 STREET: 115 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036/2711
 COMPUTER: IBM Compatible
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ Version 2.0

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/822,445
 FILING DATE: 21-MAR-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-062-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-8864
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 12225 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 FEATURE:

US-09-396-540-11

NAME/KEY: CDS
LOCATION: 190..11208

Query Match 3.2%; Score 34.4; DB 4; Length 12225;
Best Local Similarity 48.0%; Pred. No. 1.6; Indels 0; Gaps 0;

Matches 98; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 102 GGAGATAAAACAAAGGGCCTTAGAGAATTAAAAAGGTTCCCTTGACAAATACCCAGA 161
Db 8430 GGATTCAAATCAAGGATATCTAAATAGCTTGCAGATATACCCAGGTGTCTCT 8489

Qy 162 CCCGANGGAAAGAGTTAACAGGTCTGGGATTTCGGGTAAAGGAATAAGGAGAAA 221
Db 8490 CTCCCAAGGAAATGAGAAAAGGTGATCCAGCATATPAGAGGAATGTATAAGTGA 8549

Qy 222 TTAGTTCGTTACGAGACTATACCTCAATACTAGTATAGGTGA 281
Db 8550 TTGAGTGCAGCAGATTCGGAGACTATCCAGGTGACATGAGTCACTAGTGA 8609

Qy 282 ACTTACATACCCGTTACATACC 305
Db 8610 ATGGTATGCCCATCTACTATCC 8633

RESULT 9
US-09-396-540-11
Sequence 11, Application US/09396540
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA: US/09/396,540
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/822,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzz, Laura A.
; REGISTRATION NUMBER: 30,742
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..11208

Query Match 3.2%; Score 34.4; DB 2; Length 12616;
Best Local Similarity 48.0%; Pred. No. 1.6; Indels 0; Gaps 0;

Matches 98; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 102 GGAGATAAAACAAAGGGCCTTAGAGAATTAAAAAGGTTCCCTTGACAAATACCCAGA 161
Db 8430 GGATTCAAATCAAGGATATCTAAATAGCTTGCAGATATACCCAGGTGTCTCT 8489

Qy 162 CCCGANGGAAAGAGTTAACAGGTCTGGGATTTCGGGTAAAGGAATAAGGAGAAA 221
Db 8490 CTCCCAAGGAAATGAGAAAAGGTGATCCAGCATATPAGAGGAATGTATAAGTGA 8549

Qy 222 TTAGTTCGTTACGAGACTATACCTCAATACTAGTATAGGTGA 281
Db 8550 TTGAGTGCAGCAGATTCGGAGACTATCCAGGTGACATGAGTCACTAGTGA 8609

Qy 282 ACTTACATACCCGTTACATACC 305
Db 8610 ATGGTATGCCCATCTACTATCC 8633

RESULT 10
US-08-822-445-9
Sequence 9, Application US/08822445
; PATENT NO. 595223
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 435
; NAME: Coruzz, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-64
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12616 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..11592

RESULT 11
 US-09-905-173-23.rni

; Sequence 9, Application US/09396540
 ; Patent No. 6310182
 GENERAL INFORMATION:
 ; APPLICANT: Kaplan, Jerry
 ; APPLICANT: Perou, Charles
 ; APPLICANT: Moore, Karen
 TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Penne & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036/2711
 COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 2.0
 CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/396,540
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 08/822,445
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7853-062-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-790-9090
 ; TELEFAX: 212-869-8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 12616 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 190..11592
 US-09-396-540-9

QY 102 GGAGATAAAACAAAGGCCCTTAAAGAATTAAAAAGGTTCCCTTGAAACAAATACCCAGA 161
 Db 8817 GGATTCAAAATCAAAGGATATATCTAAATAGCTGAGATATCACCCAGCAGTGTCTC 8876
 QY 162 CCCGAAGGAAAGAGTTAAAGGGTTCTGGGATTTTGGCGTTAAGGAAGAAA 221
 Db 8877 CTCCCAGGAAATGAGAAAAAGGTGATCCAGCATATTAGGAAATAGTGA 8936
 QY 222 TTAGTCTGGTAACGTCGACGAACTATACCTCTAACTCTCAATAGCTATAAGTGA 281
 Db 8937 TTGAGTCAGACATGGACTTATTACAGCAGTACATAGCAGCAGTACATAGCAGT 8996
 QY 282 ACTTTACATCCGTTACATAC 305
 Db 8997 ATGGTACCCCATCTACTATCC 9020

RESULT 12
 US-08-234-939-7/c

; Sequence 7, Application US/08234939
 ; Patent No. 5459352
 GENERAL INFORMATION:
 ; APPLICANT: Conkling, Mark A.
 ; APPLICANT: Yamamoto, Yuri T.
 TITLE OF INVENTION: Root Specific Gene Promoter
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park and Gibson
 ; STREET: Post Office Drawer 34009
 ; CITY: Charlotte
 ; STATE: No. 545252th Carolina
 ; COUNTRY: U.S.A.
 ; ZIP: 28234
 COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.4
 CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/234,939
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/649,564
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sibley, Kenneth D.
 ; REGISTRATION NUMBER: 31,665
 ; REFERENCE/DOCKET NUMBER: 5051-141
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 911-881-3140
 ; TELEFAX: 919-881-3175
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1100 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)

US-08-234-939-7

Query Match 3.18; Score 33.2; DB 1; Length 1100;
 Best Local Similarity 51.3%; Pred. No. 0.88; Mismatches 73; Indels 0; Gaps 0;

QY 267 AATGCTTATGGTAACTTACATCCCTTACATCTGTCACCTTCAGATGTC 326
 Db 331 ATTGATTAAGATGTCCTCACATCCCTTACAACTATTCATCTGATGTC 272

QY 327 CGAGATAAGTGCAGAACCCCTGAAAGGTTCAACTGACGAAACT 386
 Db 271 TTGATTTGAAATTCTAAAGGAAATAGTTGCATCAAACGATTAAANGTA 212

QY 102 GGAGATAAAACAAAGGCCCTTAAAGAATTAAAAAGGTTCCCTTGAAACAAATACCCAGA 161

QY 387 TGATATAGACTTAGAAAGAAGTATTGATT 416
 Db 211 TGCCTAAATTCAAAATATTGGGGTT 182

RESULT 13
 US-08-234-939-6/c
 Sequence 7, Application US/08558865
 Patent No. 5750386

GENERAL INFORMATION:
 APPLICANT: Conkling, Mark A.
 ADDRESS: Gibson, Charles H.
 CITY: Charlotte
 STATE: No. 5750386th Carolina
 COUNTRY: U.S.A.
 ZIP: 28234

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/558,865
 FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/236,678
 FILING DATE:
 CLASSIFICATION: 800
 APPLICATION NUMBER: US/07/770,082
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Sibley, Kenneth D.
 REGISTRATION NUMBER: 31,665
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-881-3140
 TELEFAX: 919-881-3175
 TELEX: 575102

INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1100 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)

QY 387 ATATGGTAAAGTGAACATTACATACCGTTACATACCTGTCCACCTTCCATGTA 326
 Db 211 ATTGGATTAAGATGTCCTACATCCCTTACAACATTTCTACTTGTATTGTC 272

Query Match Score 33.2; DB 1; Length 1100;
 Best Local Similarity 51.3%; Pred. No. 0.88;
 Matches 77; Conservative 0; M1 smatches 73; Indels 0; Gaps 0;

QY 327 CGAGATAAGTGGAAAGAAGTCTGGAAAGACCCCTGGTAAGGGTCAACTGGACAAACT 386
 Db 439 TTAGATGTGTGATAATTCCCTAAAGGAAAATAGTGTCAAAACCATCTAAAGTA 380

RESULT 15
 US-08-558-865-6/c
 Sequence 6, Application US/08558865
 Patent No. 5750386

GENERAL INFORMATION:
 APPLICANT: Conkling, Mark A.
 ADDRESS: Gibson, Charles H.
 CITY: Charlotte
 STATE: No. 545925th Carolina
 COUNTRY: U.S.A.
 ZIP: 28234

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.24

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/234,939
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Sibley, Kenneth D.
 REGISTRATION NUMBER: 31,665
 REFERENCE/DOCKET NUMBER: 5051-141
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-881-3140
 TELEFAX: 919-881-3175
 TELEX: 575102

INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1268 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)

QY 327 CGAGATAAGTGGAAAGAAGTCTGGAAAGACCCCTGGTAAGGGTCAACTGGACAAACT 386
 Db 439 TTAGATGTGTGATAATTCCCTAAAGGAAAATAGTGTCAAAACCATCTAAAGTA 380

QY 387 TGATATAGACTTAGAAAGAAGTATTGATTGACATCAAACGATCTAAAGTA 212
 Db 327 CGAGATAAGTGGAAAGTCTGGAAAGACCCCTGGTAAGGGTCAACTGGACAAACT 386
 Db 271 TTAGATGTGTGATAATTCCCTAAAGGAAAATAGTGTCAAAACCATCTAAAGTA 212

QY 387 TGATATAGACTTAGAAAGAAGTATTGATTGACATCAAACGATCTAAAGTA 416
 Db 211 TGCCTAAATTCAAAATTTGTCATTGGGT 182

RESULT 14

ADDRESSER: Gibson
 STREET: Post Office Drawer 34009
 CITY: Charlotte
 STATE: No. 5750386th Carolina
 COUNTRY: U.S.A.
 ZIP: 28234
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/558,865
 FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/236,678
 FILING DATE:
 CLASSIFICATION: 800
 APPLICATION NUMBER: US/07/770,082
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Sibley, Kenneth D.
 REGISTRATION NUMBER: 31,665
 REFERENCE/DOCKET NUMBER: 5051-166
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-881-3140
 TELEFAX: 919-881-3175
 TELEX: 575102
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1268 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-558-865-6

Query Match 3.1%; Score 33.2; DB 1; Length 1268;
 Best Local Similarity 51.3%; Pred. No. 0.96;
 Matches 77; Conservations 0; Mismatches 73; Indels 0; Gaps 0;
 QY 267 ATAGCTATAGTGTGAACCTTACATACCCGTTACATACCTGTCACCTTCCCATGTA 326
 Db 499 ATTGGATTAAGTGAAGTGCCGCCAACCCCTTCAAACTTTCATTTGTATGTC 440
 QY 327 CGAGATACTGGAAAGTCTCGGAAGACCCCTCGTAAAGGTCAACTGGAGAAACTT 386
 Db 439 TTGATTTGTGTGATATTCTCTAAAGGAAATAGTGCATCAAACGCATCTAAGATA 380
 QY 387 TGTATAGACTAGAAAGAGTATTGAAATTGAAATTGAAATTGAAATTGAAATT 416
 Db 379 TGCCATAAAATTCAGGTTGGGTT 350

Search completed: April 15, 2003, 15:49:10
 Job time : 100 secs

Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 15:15:16 ; Search time 125 Seconds
(without alignments)

7473.469 Million cell updates/sec

Title: US-09-905-173-23

Perfect score: 1065

Sequence: atgatacccaaggattaa.....cccttcaaggctctttaa 1065

Scoring table: IDENTITY NUC Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/us07_pubcomb.seq;*

2: /cgn2_6/ptodata/2/pubpna/pct_new_pub_seq;*

3: /cgn2_6/ptodata/2/pubpna/us06_new_pub_seq;*

4: /cgn2_6/ptodata/2/pubpna/us06_pubcomb.seq;*

5: /cgn2_6/ptodata/2/pubpna/us07_new_pub_seq;*

6: /cgn2_6/ptodata/2/pubpna/pctus_new_pubcomb.seq;*

7: /cgn2_6/ptodata/2/pubpna/us08_new_pub_seq;*

8: /cgn2_6/ptodata/2/pubpna/us08_pubcomb.seq;*

9: /cgn2_6/ptodata/2/pubpna/us09_new_pub_seq;*

10: /cgn2_6/ptodata/2/pubpna/us09_pubcomb.seq;*

11: /cgn2_6/ptodata/2/pubpna/us10_new_pub_seq;*

12: /cgn2_6/ptodata/2/pubpna/us10_pubcomb.seq;*

13: /cgn2_6/ptodata/2/pubpna/us60_new_pub_seq;*

14: /cgn2_6/ptodata/2/pubpna/us60_pubcomb.seq;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1065	100.0	1065	9	US-10-060-432-23		Sequence 23, Appl
2	1065	100.0	1065	10	US-09-905-173-23		Sequence 23, Appl
3	46.2	4.3	1085	10	US-09-974-300-298		Sequence 298, Appl
4	42.2	4.0	1092	10	US-09-905-173-35		Sequence 579, Appl
5	37.8	3.5	198	9	US-10-083-357-579		Sequence 579, Appl
6	37.6	3.5	588	10	US-09-864-761-6646		Sequence 6646, Appl
7	36.2	3.4	6804	7	US-08-781-986A-365		Sequence 365, Appl
8	35.8	3.4	560	10	US-09-94-300-4716		Sequence 4716, Appl
9	35.8	3.4	1098	9	US-09-738-626-2298		Sequence 2298, Appl
10	35.2	3.3	1350	9	US-09-971-536-8		Sequence 8, Appl
c 11	34.6	3.2	2000	9	US-09-938-842A-3936		Sequence 3336, Appl
c 12	34.4	3.2	375	10	US-09-864-761-1114		Sequence 11114, Appl
c 13	34.4	3.2	2270	10	US-09-887-576-455		Sequence 455, Appl
c 14	34.4	3.2	12225	10	US-09-927-668-11		Sequence 11, Appl
c 15	34.4	3.2	12616	10	US-09-927-668-9		Sequence 9, Appl
c 16	34.2	3.2	534	9	US-09-738-626-2011		Sequence 2011, Appl
c 17	34.2	3.2	1093	9	US-10-001-876-78		Sequence 78, Appl
c 18	34.2	3.2	1691139	9	US-10-067-514-1		Sequence 1, Appl
c 19	34.2	3.2	3309400	9	US-09-738-626-28		Sequence 5128, Appl

ALIGNMENTS

Result No.	Score	Query	Match	Length	DB	ID	Description
1	671	9	US-10-184-644-346				Sequence 346, Appl
2	671	9	US-09-974-300-2020				Sequence 2020, Appl
21	33.8	3.2	1851	10	US-09-738-626-245		Sequence 245, Appl
22	33.6	3.2	1023	10	US-09-948-649-1		Sequence 1, Appl
23	33.2	3.1	1480	10	US-09-974-300-7726		Sequence 7726, Appl
24	33	3.1	837	10	US-09-974-300-7714		Sequence 7714, Appl
25	33	3.1	948	10	US-09-938-842A-4699		Sequence 4699, Appl
26	32.8	3.1	1757	9	US-09-974-300-1427		Sequence 1427, Appl
27	32.6	3.1	319	10	US-09-974-300-1427		Sequence 1427, Appl
28	32.6	3.1	446	10	US-09-864-761-1815		Sequence 3, Appl
c 29	32.6	3.1	53332	9	US-10-224-562-3		Sequence 3, Appl
c 30	32.6	3.1	53332	10	US-09-801-861-3		Sequence 3, Appl
c 31	32.4	3.0	668	10	US-08-781-986A-309		Sequence 98, Appl
c 32	32.4	3.0	1872	10	US-09-884-761-24336		Sequence 23380, Appl
c 33	32.4	3.0	1884	10	US-09-815-242-8207		Sequence 8207, Appl
c 34	32.2	3.0	477	12	US-10-038-780-12		Sequence 449, Appl
c 35	32	3.0	336	10	US-09-969-708-449		Sequence 312, Appl
c 36	32	3.0	336	10	US-09-347-312		Sequence 312, Appl
c 37	31.8	3.0	472	9	US-10-184-644-98		Sequence 157, Appl
c 38	31.6	3.0	161	10	US-09-867-761-157		Sequence 66, Appl
c 39	31.6	3.0	5134	10	US-09-757-764-903-66		Sequence 100, Appl
c 40	31.6	3.0	32768	10	US-09-070-92-A-100		Sequence 3, Appl
c 41	31.6	3.0	32816	10	US-09-329-094-3		Sequence 21849, Appl
c 42	31.4	2.9	242	10	US-09-864-761-21849		Sequence 160, Appl
c 43	31.4	2.9	392	9	US-10-184-644-160		Sequence 5128, Appl
c 44	31.4	2.9	498	10	US-09-864-761-5128		Sequence 5128, Appl

RESULT 1

; Sequence 23, Application US-10-060-432-23

; Publication No. US200304092A1

; GENERAL INFORMATION:

; APPLICANT: WARREN, Patrick V.

; SWANSON, Ronald V.

; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

; CECCHI, STEWART & OLSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US-10-060-432

FILING DATE: 29-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/481,733

FILING DATE: 11-JAN-2000

APPLICATION NUMBER: US-08/599,171

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: HERRON, CHARLES J.

REGISTRATION NUMBER: 28,019

REFERENCE/DOCKET NUMBER: 331400-38

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 1065 NUCLEOTIDES

TYPE: NUCLEIC ACID

Query Match 4.3%; Score 46.2; DB 10; Length 1085;
 Best Local Similarity 47.4%; P ref. No. 0.00077;
 Matches 138; Conservative 0; Mismatches 153;
 Indels 0; Gaps 0;

Qy 361 GTAAAGGTCAACTGGACGAAACCTTGTATAGACTTAAAGAGTATTGAATTATA 420
 Db 361 GTAAAGGTCAACTGGACGAAACCTTGTATAGACTTAAAGAGTATTGAATTATA 420

Qy 421 GAGAAAGAAAACCGTTCTGGTACTTGCTTACCAACAAACCCACGGAAACCTC 480
 Db 421 GAGAAAGAAAACCGTTCTGGPACTTGCTTACCCAAACACCCACGGAAACCTC 480

Qy 481 TTTCCAGGGAAAGATTGAGGATAAGAAACAGGGTTCTGCTAATAGACGAA 540
 Db 481 TTTCCAGGGAAAGATTGAGGATAAGAAACAGGGTTCTGCTAATAGACGAA 540

Qy 541 GCCTACTATCATTCATTCACCGGAAACCTTCTGGAGACGCGCTCAAAGGGAAAGATAG 600
 Db 541 GCCTACTATCATTCACCGGAAACCTTCTGGAGACGCGCTCAAAGGGAAAGATAG 600

Qy 601 GTAGTTTGGAGGACACTTCAAAATCGTATGGGAGTTAAAGGTAGGGATTAA 660
 Db 601 GTAGTTTGGAGGACACTTCAAAATCGTATGGGAGCTTAAAGGTAGGGATTAA 660

Qy 661 GGAAGGGGAAATCTCTCAGAAATTAAACAGGTGAGACTCCCTCACGTGACCTAC 720
 Db 661 GGAAGGGGAAATCTCTCAGAAATTAAACAGGTGAGACTCCCTCACGTGACCTAC 720

Qy 721 CCCTCTAGGTGATGGCAAAAGTTCTCCATCGGAGGAGAAATTCTCTAATGGAAAAAG 780
 Db 721 CCCTCTAGGTGATGGCAAAAGTTCTCCATCGGAGGAGAAATTCTCTAATGGAAAAAG 780

RESULT 4
 US-09-905-173-35

; Sequence 35, Application US/09905173
 ; Patent No. US2002013295A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; ATTORNEY: SHORT, JAY M.
 ; APPLICANT: WARREN, PATRICK V.
 ; APPLICANT: SWANSON, RONALD V.
 ; APPLICANT: MATHUR, ERIC J.
 ; TITLE OF INVENTION: ENZYMES HAVING TRANSAMINASE AND AMINOTRANSFERASE ACTIVITY AND
 ; TITLE OF INVENTION: USE THEREOF
 ; CURRENT APPLICATION NUMBER: US/09/905_173
 ; CURRENT FILING DATE: 2001-07-12
 ; PRIOR APPLICATION NUMBER: US 09/412,184
 ; PRIOR FILING DATE: 1999-10-04
 ; PRIOR APPLICATION NUMBER: US 09/389,537
 ; PRIOR FILING DATE: 1999-09-02
 ; PRIOR APPLICATION NUMBER: DIVER11240-7
 ; PRIOR FILING DATE: 1996-05-08
 ; PRIOR APPLICATION NUMBER: US 08/599,171
 ; PRIOR FILING DATE: 1996-02-09
 ; PRIOR APPLICATION NUMBER: US 09/481,733
 ; PRIOR FILING DATE: 2000-01-11
 ; PRIOR APPLICATION NUMBER: US 09/069,226
 ; ORGANISM: Ammonifex degensii
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (986). (987)
 ; OTHER INFORMATION: n is any nucleotide
 ; SEQ ID NO: 35
 ; LENGTH: 1092
 ; TYPE: DNA

Query Match 4.0%; Score 42.2; DB 10; Length 1092;
 Best Local Similarity 48.2%; P ref. No. 0.013;
 Matches 119; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 138 GGTCCTTGACAAATCCAGACCCAGAAAGGTAAAGCGTACGGCTCGGGAA 197
 Db 168 GGACCAATGGCACCTTACCCAGAAAGCTCACGTGACGGCTGGTAA 227

Qy 198 TTTTTCGGCTTAAGGAAGAAAATTAGTTCCTCGGTAAACGGTTGGACGAACCTATATA 257

RESULT 3
 US-09-974-300-298

; Sequence 298, Application US/09974300
 ; GENERAL INFORMATION:
 ; APPLICANT: Berka, Randy M.
 ; APPLICANT: Clausen, Ib Groth
 ; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
 ; TITLE OF INVENTION: Expression
 ; FILE REFERENCE: 10085_500-US
 ; CURRENT APPLICATION NUMBER: US/09/974,300
 ; PRIOR APPLICATION NUMBER: 09/680,598
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR APPLICATION NUMBER: 60/279,526
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 8481
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 298
 ; LENGTH: 1085
 ; TYPE: DNA
 ; ORGANISM: Bacillus licheniformis
 ; US-09-974-300-298

CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,986A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Benson, Bob
 REGISTRATION NUMBER: 30,446
 REFERENCE/DOCKET NUMBER: PB248PP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 365:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6804 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 ;
 US-08-781-986A-365

Query Match 3.4%; Score 36.2; DB 7; Length 6804;
 Best Local Similarity 57.5%; Pred. No. 2.7; Indels 0; Gaps 0;
 Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 337 GCGAAAGTCTCGGAAGAACCCCTCGTAAAGGTCAACTGGACAAAATTGATATGAC 396
 Db 644 GCTCAGATTTAAAGTCGGCATGAAACTGATTAATGCCATTATGATTAGAT 703

Qy 397 TTGAAAGAAGTATGAAATAAGAGAAAACCGTTCGGTACPT 449
 Db 704 GAAGGAAACGTACTGAAATGGAAAAATAATGCTATGAGGGTCAT 756

RESULT 8
 US-09-974-300-4716
 Sequence 4716, Application US/09974300
 Patent No. US20020146721A1
 GENERAL INFORMATION:
 APPLICANT: Berka, Randy M.
 ATTORNEY: Clausen, Ib Groth
 TITLE OF INVENTION: Methods For Monitoring Multiple Gene
 TITLE OF INVENTION: Expression
 FILE REFERENCE: 10085, 500 US
 CURRENT APPLICATION NUMBER: US/09/974,300
 CURRENT FILING DATE: 2001-10-05
 PRIOR APPLICATION NUMBER: 09/680,598
 PRIOR FILING DATE: 2000-10-06
 PRIOR APPLICATION NUMBER: 60/279,526
 PRIOR FILING DATE: 2001-03-27
 NUMBER OF SEQ ID NOS: 8481
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 4716
 LENGTH: 560
 TYPE: DNA
 ORGANISM: Bacillus clausii
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(560)
 OTHER INFORMATION: n = A,T,C or G

US-09-974-300-4716

Query Match 3.4%; Score 35.8; DB 10; Length 560;

Best Local Similarity 45.1%; Pred. No. 0.84;
 Matches 133; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

Qy 154 TACCCAGACCCCCAAGCAGAAAGTAAAGGGTCTTGCGGATTTTGGCGTTAAG 213
 Db 181 TAATCGGTATGGCTATGGACGCTTGCCAAAGCATAAAGTAGAT 240

Qy 214 GAAGAAAATTAGTTAGTCTCGGTAAAGGTICGGACAACTCATATACCTCTCAATAGCT 273
 Db 241 GAAGGACAGCTTATAGGAAATGGACCGATGAATTTACATGTAAGGAGT 300

Qy 274 ATAGGTGAACTTACATACCCGTTACATACCTGTCACCTTCCATGTAAGGAGT 333
 Db 301 TTTTAAGGCTGAGACCAAACGGTTATGGGACGCCCTTTCTCAGTACAAGTTA 360

Qy 334 AGTGCGAAAGTCTCGGAAAGACCCCTCGTAAAGGTAAACTGGACAAAATTGATATA 393
 Db 361 AAATGCAAGATGAGGGCAACGATTAAGAAATAACCGTTAAAGAACGGCTCCAC 420

Qy 394 GACTTAGAAGAAGTATTGAAATAATGAGAAAGAAAACCGTTCTGGTACT 448
 Db 421 GATTAGATGCATGGCTTGACGCCAAACGGGATTGACGCCAAACGGGATTCTGGGTCT 475

RESULT 9
 US-09-738-626-2298
 Sequence 2298, Application US/09738626
 Publication No. US20020137605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 NUMBER OF SEQ ID NOS: 7059
 SEQ ID NO 2298
 LENGTH: 1098
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-2298

Query Match 3.4%; Score 35.8; DB 9; Length 1098;
 Best Local Similarity 45.5%; Pred. No. 1.2; Indels 0; Gaps 0;

Qy 235 AACGGTTGGGACGAACTCATACTCATCTCAATAGCTATAGGTGAACTTTACATACCC 294
 Db 286 AATGGTCCAAATGAAATTCTGAGCTGCTGGCTGACCTGGCAGC 345

Qy 295 GTTACATACCTGTCGCCACCTTCCATGAGATAAGTGGAAACTTCCTGGAGA 354
 Db 346 GCCTGGGATTCACGCCAGCTATTCATCACCCATTTGGCTAAAGGCCACCACT 405

Qy 355 CCCCTGGTAAAGGTAAACTGGACAAAATTGATGACTAGAAAAGAAACTATGAA 414
 Db 406 GAATTCAATTGGGTGCCCAGGTCTGATTCCGCATATGATGGCTGGCTGGAA 465

Qy 415 TTAATAGAGAAAGAAAACCCGTTCTGGGTACTTGTACCCAAACCCCCACGGGA 474
 Db 466 GAAATCGTGAAGAGCCGTACATGTTGTCACCCGAACCCGACCGGT 525

Qy 475 ACCCTTTTCAGGGAAAGATGGAGATAAGAAC 513
 Db 526 GATGTGACCTCGTGCAGATGTTAGCGCATCACAA 564

RESULT 10
 US-09-971-536-B
 ; Sequence 8, Application US/09971536
 ; Patent No. US2002015997E1
 ; GENERAL INFORMATION:
 ; APPLICANT: Glenn, Matthew
 ; Haukkala, Tikka
 ; Bloksberg, Leonard
 ; Lubbers, Mark
 ; Dekker, James
 ; Christansson, Anna
 ; APPLICANT: Holland, Ross
 ; O'Toole, Paul
 ; Reid, Julian
 ; APPLICANT: Coolbear, Timothy
 ; TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Methods
 ; TITLE OF INVENTION: Using Them
 ; FILE REFERENCE: 1043c2
 ; CURRENT APPLICATION NUMBER: US/09/971.536
 ; CURRENT FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: U.S. No. US2002015997E1 09/634,238
 ; PRIOR FILING DATE: 2000-08-08
 ; PRIOR APPLICATION NUMBER: U.S. No. US2002015997E1 09/724,623
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: PCT/NZ01/00160
 ; PRIOR FILING DATE: 2001-08-08
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 1350
 ; TYPE: DNA
 ; ORGANISM: Lactobacillus rhamnosus
 US-09-971-536-B

Query Match 3.38; Score 35.2; DB 9; Length 1350;
 Best Local Similarity 46.1%; Pred. No. 2.1; Indels 3; Gaps 1;
 Matches 155; Conservative

Qy 152 ATATCCAGACCCGAGCAGAAAGAGTTAACGGGTTCTCGGATTTTCGGCTTA 211
 Db 304 ATTACCCGATGCGCTATGCACTCAACTGGCAAGTGTAGCGTTG 363

Qy 212 AGGAGAAAATTAGTTGGTACCGTGGACATATACTCTAATAG 271
 Db 364 CGCGAGACAGTGGTTTGCAAGTGTAGCAAGTGTAGCGTTG 423

Qy 272 CTATAGGTGAACTTACATACCGTTACATACCTGTCACCTTCAGTACGAGA 331
 Db 424 CTTTTAGCCGGGGATGGAAGTCAATGCCAACATGGCGT 483

Qy 332 TAACTGGAAAGTCTGGAAACCCCTCGTAAAGGTCAACTGGACAAACTTGATA 391
 Db 484 TGCAATGCCAAATGAAAGGGGCCACCTGATGTCGCTCACTGGAAATT 543

Qy 392 TAGACTTAGAAGAGATTTGAAAGAAACCCGTTCTCGGAAACCTTGATA 451
 Db 544 TTGATTATCTGAATGCCAGGGCGCTACTGAAACTGGTGT 600

Qy 452 CCTACCCAAAACCCCCACGGAAACCTCTTCCCA 487
 Db 601 GCAACCAAATAACCCACGGGACGCGTGTCAA 636

us-09-938-842A-3936/C
 ; Sequence 3936, Application US/09938842A
 ; Patent No. US2002010378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; Keeps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIP1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIORITY NUMBER: US 60/227,866
 ; PRIORITY FILING DATE: 2000-08-24
 ; PRIORITY APPLICATION NUMBER: US 60/264,647
 ; PRIORITY FILING DATE: 2001-01-16
 ; PRIORITY APPLICATION NUMBER: US 60/300,111
 ; PRIORITY FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 3936
 ; LENGTH: 2000
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-3936

Query Match 3.28; Score 34.6; DB 9; Length 2000;
 Best Local Similarity 59.8%; Pred. No. 4.1; Indels 0; Gaps 0;
 Matches 58; Conservative 0; Mismatches 39;

Qy 822 AATGAAGAAATAAGAAGGTTGAGGTTTCGGAGTAACTTGCTTTTCAG 881
 Db 1625 AAAACAGAGTTGAGAAGTTGCAAGATCTCCAAATAAGGTCAGTGTATTCAC 1566

Qy 882 AACGCCTTACCCGCCACAGGGTTATCAGGAGCTA 918
 Db 1565 AAGTTTTAATCIGCTTAATAGGTTATGATCATCTA 1529

RESULT 12
 US-09-864-761-11114
 ; Sequence 11114, Application US/09864761
 ; Patent No. US2002004876A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; Hanzel, David R.
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Hanzel, David R.
 ; APPLICANT: Chen, Weisheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aeomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIORITY NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIORITY NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIORITY NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIORITY NUMBER: GB 24263.6
 ; PRIORITY NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIORITY NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIORITY NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIORITY NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIORITY NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIORITY NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIORITY NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006660
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-10
 PRIOR APPLICATION NUMBER: US 09//774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annotrax Sequence Listing Engine vers. 1.1
 SEQ ID NO 11114
 LENGTH: 375
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC006212.1
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.65
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.58
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.77
 US-09-864-761-11114

Query Match 3.2%; Score 34.4; DB 10; Length 2270;
 Best Local Similarity 54.8%; Pred. No. 5.1;
 Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 738 AAAAGTTCCTCCACGGAGGAACTTCATAATGCCAAGATACAGGAGTTGTAC 797
 Db 1680 AGATTTACCATATGAGATCATGATTAATGCTGAGGAGATATTGAGATGAAAC 1621

QY 798 AGAGCGAGAAAGGATGTACGACCAATGAGAAAATAGAGGTTGAGTTTTCGGAG 857
 Db 1620 AGACAAATAAGGTATCATTCGCCCAATGAGAACATAACAGTTAGAACAGTGGCCCCGA 1561

QY 858 TAAG 861
 Db 1560 TAAG 1557

RESULT 14
 US-09-927-668-11
 Sequence 11, Application US/09927668
 Patent No. US200211514A1
 GENERAL INFORMATION:
 APPLICANT: Kaplan, Jerry
 PEROU, Charles
 Moore, Karen
 TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESS: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036/2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/927,668
 FILING DATE: 10-Aug-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/396,540
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-062-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-859-8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 12225 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 190...11208
 SEQUENCE DESCRIPTION: SEQ ID NO: 11:
 US-09-927-668-11

Query Match 3.2%; Score 34.4; DB 10; Length 12225;

Result No.	Score	Query Match Length	DB ID	Description
1	65.8	6.2	573 10 BE432349	BE432349 EST398878 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone CLEG7J1, mRNA sequence.
2	65	6.1	587 10 AW034522	AW034522 EST398878 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone CLEG7J1, mRNA sequence.
3	62.6	5.9	819 12 BG35830	BG35830 09FG11 Ma
4	58.8	5.5	762 13 BM412207	BM412207 EST586534
5	58.4	5.5	631 13 BJ463860	BJ463860 EST518029
6	57.2	5.4	778 12 BG887294	BG887294 EST513145

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT	JOURNAL	COMMENT
BE432349	Unpublished (2000)	Clemson University Genomics Institute
BE432349	Unpublished (2000)	Clemson University
BE432349	Unpublished (2000)	100 Jordan Hall, Clemson, SC 29631, USA
BE432349	Unpublished (2000)	Email: http://www.genome.clemson.edu/orders/index.html
BE432349	Unpublished (2000)	5 prime sequence.
BE432349	Unpublished (2000)	Location/Qualifiers
BE432349	Unpublished (2000)	1 . 573

ALIGNMENTS

LOCUS	DEFINITION	ACCESSION	VERSION	ORGANISM	KEYWORDS	SOURCE	REFERENCE
EST398878	tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone CLEG7J1, mRNA sequence.	BE432349	BE432349	Lycopersicon esculentum	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; Eudicots; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.	tomato.	Arcalda,J., Vrebalov,J.J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
AW034522		BE432349	BE432349				
BG35830		BE432349	BE432349				
BM412207		BE432349	BE432349				
BJ463860		BE432349	BE432349				
BG887294		BE432349	BE432349				

LycoPersicon esculentum														
Organism="LycoPersicon esculentum"														
/cultivar="T496"														
/db_xref="Taxon:4081"														
/clone="CLEC24MB"														
/clone_lib="tomato breaker fruit, TIGR"														
/feature_type="Pericarp"														
/dev_stage="breaker"														
/lab_host="SOLR"														
/note="Vector: pBluescriptSKmCUadapt; Site_1: ECOR1; Site_2: CLEC24MB"														
Site_1: XbaI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."														
156 a 120 c 128 g 169 t														
FEATURES	source	1.	587	Location/Qualifiers										
		/organism="LycoPersicon esculentum"												
		/cultivar="T496"												
		/db_xref="taxon:4081"												
		/clone="CLEC24MB"												
		/clone_lib="tomato callus, TAMU"												
		/feature_type="callus"												
		/dev_stage="25-40 days old"												
		/lab_host="XL1-Blue MRF"												
		/note="Vector: pBluescript SK(-); Site_1: ECOR1; Site_2: XbaI; supplier: Giovanni laboratory; cLEC - Colecidons of seedlings 7-10 days post-germination were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library."												
		5 prime sequence.												
		Location/Qualifiers												
		1. 587												
BASE COUNT	ORIGIN	104	c	148	g	172	t							
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 3.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 193; Conservative 0; Mismatches 212; Indels 0; Gaps 0;	Similarity	47.1%												
	Matches	200;	Indels	0;	Gaps	0;								
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225;保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								
Best Local Similarity Pred. No. 5.1e-07;	Local	47.1%		Pred. No. 5.1e-07;										
Matches 225; 保守性 0; 错配 225; 插入 0; 缺失 0;	相似度	47.1%												
Query Match Score 65.8; DB 10; Length 587;	Match	6.1%		Score 65;	DB 10;	Length 587;								

AUTHORS	Nielsen,K.L., Crookshanks,M., Emmersen,J. and Welinder,K.G.
JOURNAL	EST-sequencing of mature potato tuber (Var. Kuras)
COMMENT	Unpublished (2000)
Contact	Karen G. Welinder Institut for bioteknologi Aalborg Universitet Sohngardsholmvej 49, 9000 Aalborg, Denmark
TELE	+45 96358467
FAX	+45 98141808
Email	kgw@bio.auc.dk
Sequenced from the 5' end.	
High quality sequence stop	819
POLY-A-No.	
FEATURES	Location/Qualifiers
source	1..819 /organism="Solanum tuberosum" /cultivar="Field grown Kuras" /db_xref="Field:4113" /clone_lib="Mature tuber lambda ZAP" /tissue_type="Tuber" /note="Vector: Lambda ZAP"
BASE COUNT	227 a 155 c 194 g 243 t
ORIGIN	
Query Match	5.9%; Score 62.6; DB 12; Length 819;
Best Local Similarity	47.0%; Pred. No. 2.6e-06;
Matches 194; Conservative	0; Mismatches 219; Indels 0; Gaps 0;
Qy	149 ACAAAATACCCAGACGCCGAAAGCGTTAACGGGTTCTGGATTTCGGC 208
Db	393 ATATTTATCCCTGATCCTGAAGGCCGTACCTGCGCACTCTGCTGAGATCTGGC 452
Qy	209 TAAAGGAAAGAAAATTAGTCTCGTAACTGGTCCGACACTATATACTCTCTAA 268
Db	453 TTGAATCTGGATGATATTCTTGAGCTGGTGGAGATGAACCTATTGTTGATAATGA 512
Qy	269 TAGCTATAGGTGAACTTATACCCGTTACATACCTGTTCCCACTATGTA 328
Db	513 GATGGCTATGGATCTGGAGCAGATGTTGGACTGCCACCCACTTACAATGTTG 572
Qy	329 AGATAAGTGCAGAAAGTCTCGGAAGAACCCCTCGTAAGGGTTCACTGGAGAAACTTGT 388
Db	573 AATTGATGGCAGCTGTAATGGAGCAAGCTCATCAAGGTGCCCTAGGAACCCAGACTTA 632
Qy	389 ATATAGACTTAAAGAAGATTGATTGATTGATTATAGAAAGAAAAACCGTTCTCGGTACT 448
Db	633 GCCTGGATGATGAAACGGATGTCGGAAAGTGGTGAACCTGAGAACCGAGATATTTC 692
Qy	449 TTGGCTTACCCAAACACCACCCAGGGAAACCTCTTCCAGGGAAAGATTGAGGAGATA 508
Db	693 TGACATCACCTAAATACCTGATGGAGTAAATGGATGAAACTTATGAAATAC 752
Qy	509 GAAACAGGGTGTCTGTGTAATAGACGAAGGCTACATCATCTCGGA 561
Db	753 TTGACCTGCCCATATTGGATAATGGATGAAAGCATATCTGATGGGA 805
RESULT 4	
LOCUS	BM412207
DEFINITION	ESTS86534 tomato breaker fruit Lycopersicon esculentum cDNA clone cLEG59E12 5' end. mRNA sequence.
ACCESSION	BM412207
KEYWORDS	EST.
SOURCE	tomato
ORGANISM	Lycopersicon esculentum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon; Solanum; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
REFERENCE	1 (bases 1 to 767)
AUTHORS	Alcalá,J., Vrebalov,J., White,R., Vision,T., Karanycheva,S.A., Tsai

TITLE	J., Bougrí, O., Kirkness, E., Utterback, T., Van Aken, S., Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
JOURNAL	Generation of ESTs from tomato fruit tissue, breaker stage (2002)
COMMENT	Unpublished (2002)
Contact	CIGI Clemson University Genomics Institute Clemson University
TELE	100 Jordan Hall, Clemson, SC 29634, USA
FAX	Email: http://www.genome.clemson.edu/orders/index.html
Email	This clone is available through the Clemson University Genomics Institute
Seq Primer	T ³ .
FEATURES	Location/Qualifiers
source	1..762 /organism="Lycopersicon esculentum" /cultivar="TA496" /db_xref="Taxon:081" /clone="cLEG59E12" /clone_lib="tomato breaker fruit" /tissue_type="Pericarp" /dev_stage="breaker" /lab_host="SOLR" /note="Vector: phluerscriptSKmcuadapt; Site_1: EcorI;"
site	Site 2: Xhol; supplier: Boyce Thompson Institute; Fruit sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp.
BASE COUNT	207 a 155 c 178 g 222 t
ORIGIN	
Query Match	5.5%; Score 58.8%; Pred. No. 2.9e-05;
Best Local Similarity	48.8%;保守性 0; Mismatches 167; Indels 0; Gaps 0;
Matches 159; Conservative	
Qy	149 ACAAAATACCCAGACCCGAAAGCGTTAACGGGTTCTGGGATTTTCGGCG 208
Db	395 ACATTATCCTGCTGAAAGCCGACGGCTGGTGAAGCTGATCTGCG 454
Qy	209 TTAAGGAAAGAAAATTAGTCTCGGAAAGCTGAGCTATATACTACCTCTCAA 268
Db	455 TTGAATCTGAGTATCTGAGGGTGGGNGAGATGACTCATGTGATATGA 514
Qy	269 TAGTATAGGTGAACTTTACATACCCGTTACATACCTGTTCCCATGTACG 328
Db	515 GATGCTATTTGGATCTGGTCAAGGTTGGACTGCCACCTTCACATGTATG 574
Qy	329 AGATAAGTGCAGAAAGTCTCGGAAGAACCCCTCGTAAGGGTCAACTGGACAAACTTG 388
Db	575 AATTGATGGCTTAATGGAGCTGATCTGGCTGCTGAGATTTATGATATG 634
Qy	389 ATATAGACTTAAAGAAGATTGATTGATTGATTATAGAAAGAAAAACCGTTCTCGGTACT 448
Db	635 GCCTGGATGTTGAACGGTATTACCGGAACTGGCTGAGAACGGTCAAGTGGCTAGAAACCCAGACTTA 694
Qy	449 TTCTCTACCCAAACAAACCCAGGGAAAGTGGAGTAAATGGATGAAACTTATGAAATAC 720
Db	695 GACATCACCTAAATCTGTATGGGA 474
RESULT 5	
LOCUS	BM412207
DEFINITION	ESTS86534 tomato breaker fruit Lycopersicon esculentum cDNA clone cLEG59E12 5' end. mRNA sequence.
ACCESSION	BM412207
KEYWORDS	EST.
SOURCE	tomato
ORGANISM	Lycopersicon esculentum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; euasterids I; Solanales; Solanaceae; Solanum;
REFERENCE	1 (bases 1 to 767)
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; euasterids I; Solanales; Solanaceae; Solanum;
TITLE	LINEAR EST 23-MAY-2002
JOURNAL	mRNA linear
COMMENT	631 bp mRNA library, cv. Haruna Nijo germination shoots Hordeum vulgare subsp. vulgare cdNA clone bags30m21 5', mRNA sequence.
DEFINITION	BJ463860
LOCUS	BJ463860 K. Sato unpublished CDNA library, cv. Haruna Nijo
VERSION	EST.
KEYWORDS	EST.
SOURCE	Hordeum vulgare subsp. vulgare
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS	van der Hoeven, R., Bezzericides, J., Sun, H., Cho, J., Chiemingo, A., Bougri, O., Bueil, C.R., Ronning, C., Tanksley, S. and Baker, B.
TITLE	Generations of ESTs from sprouting potato eyes
JOURNAL	Unpublished (2000)
COMMENT	Contact: Cathy Ronning The Institute for Genomic Research For clone info: Please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@ressgen.com Seq primer: M13RF.R.
FEATURES	<p>Location/Qualifiers</p> <p>1. .811 /organism="Solanum tuberosum" /cultivar="Rennebec" /db_xref="taxon:4113" /clone="cST22C22" /clone_lib="cSTS" /tissue_type="sprouting eyes from tubers" /dev_stage="12-14 weeks post harvest" /lab_host="SOLR", note="Vector: pBluescript SK(); Site_1: EcorI; Site_2: XbaI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers." 227 a 144 c 198 g 242 t</p>
BASE COUNT	
ORIGIN	
Source	<p>Query Match Best Local Similarity 5.0%; Score 53.4; DB 12; Length 811; Matches 159; Conservative 0; Mismatches 176; Indels 0; Gaps 0;</p> <p>Qy 149 ACAATAPCCAGACCCGAAAGCGTAAAGAGT 5.7%; Pred. No. 0.00091; Site_1: EcorI; Site_2: XbaI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."</p> <p>Db 368 ATATTATCCCTGATCTGTGACTATTTCTGGAAAGCCGTAGCTTGAGCTCTTGCTGAAGATTCCTGCC 4.27</p> <p>Qy 209 TTAAAGAAGAAAATTATGTTCTCGGTAAACGGTTCTGGGAACACTCATATACTACCTCTCAA 2.68</p> <p>Db 428 TTGAATCTGACTATATTCTGGGGTGGCAGTAACCTGATTGATATGTA 4.87</p> <p>Qy 269 TAGCTATAGGTGAACTTTACATACCCGGTTACATACTGTGTCCTGCACCTTCCATGACG 3.28</p> <p>Db 488 GATGGTAGTGGATCTGGTACAATGGTGAACCCCACCTTCACAATGTATG 5.47</p> <p>Qy 329 AGATAAGTGGAAAGATCTGGAAACCCCTGTAAGAGTTCAACTTGAGAAACCTTGT 3.88</p> <p>Db 548 AAATGTGATGGCTGAAATGGAGCTGATCAAGGGTGCCTAGGAACCCAGACTTTA 6.07</p> <p>Qy 389 ATATAGACTTGAAGAACATATTGATTAATGAGAAAGAAAACCC 4.48</p> <p>Db 608 GCCTGTGATGTTAGAACGGATGCCGAAAGTGTTGAACTGTGAAACCGAAGTGATTT 6.67</p> <p>Qy 449 TTGCTTACCCAAACACCCACGGAAACCTCTT 4.83</p> <p>Db 668 TGACATCACCTAAATATCCTGAGGAGTATTT 7.02</p>
RESULT	8
LOCUS	BM322300
DEFINITION	PC1.2_G06..b1_A002 Pathogen infected compatible 1 (P1C1) Sorghum bicolor CDNA, mRNA sequence.
ACCESSION	BN322300
VERSION	BN322300.1 GI:18059151
KEYWORDS	
ORGANISM	Sorghum bicolor Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 622)
REFERENCE	Cordonnier-Pratt, M.-M., Gingle, A., Fang, G.C., Dean, R., Wing, R., Sudman, M., and Pratt, J.H.
AUTHORS	

RESULT	9	Qy	484	TCCAGGGAAAGATTGAGGATAAGAACAGGGTTTCTGTAAATAGACGGGC 543	
BM812973	639 bp mRNA linear EST 05-MAR-2002	Db	313	GACCATGAGATCTTCAGATCTGAAGATCTGCTCCCATCTGGCGCTGATGAGCG 372	
LOCUS	EST591066 BNIR Medicago truncatula/Meloidogyne incognita mixed EST library cDNA clone PBNIR-4N3, mRNA sequence.	DEFINITION	Qy	544	TACTATCATTACTC 557
ACCESSION	BM812973	VERSION	Db	373	TACATTGAGTTTC 386
KEYWORDS	EST	SOURCE	RESULT	10	
ORGANISM	Medicago truncatula/Meloidogyne incognita mixed EST library.	LOCUS	BJ280409	555 bp mRNA linear EST 09-APR-2002	
Eukaryota; mixed EST libraries.	REFERENCE	DEFINITION	BJ280409	Y. Ogihara unpublished cDNA library, wh_r Triticum aestivum cDNA clone whr7p24 5', mRNA sequence.	
AUTHORS	Bird,D., Koltai,H., Samac,D., Town,C.D., Van Aken,S., Utterback,T., Cheung,F., Tsai,J. and Fraser,C.M.	ACCESSION	BJ280409		
TITLE	ESTs from roots of Medicago truncatula after infection with the nematode Meloidogyne incognita	VERSTION	BJ280409.1	GI:2102528	
COMMENT	Unpublished (2002)	KEYWORDS	EST.		
CONTACT:	Bird, DM	ORGANISM	bread wheat.		
Plant Nematode Genetics Group	Box 7616, Raleigh, NC 27695, USA	LOCATION	Triticum aestivum		
Tel: 919 515 6813	Fax: 919 515 9500	QUALIFIERS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae		
E-mail: david_bird@ncsu.edu	TIGR sequence name: MTQAC74TK	REFERENCE	Triticeae; Triticum.		
www.medicago.org	More information is available at: 1. (bases 1 to 555)	AUTHORS	Ogihara,Y. and Murai,K.		
Seq primer: SKm02 (CTA GAA CTA gtg gAT CC).	/organism="Medicago truncatula/Meloidogyne incognita mixed	TITLE	Expressed genes in Triticum aestivum		
Location/Qualifiers	incognita" second stage larvae"	JOURNAL	Unpublished (2002)		
1. 639	/dev_stage="3 week old roots"	COMMENT	Contact: Tadatsu Shin-i		
/organism="Medicago truncatula/Meloidogyne incognita mixed	/dev_stage="3 days after infection with Meloidogyne	CONTACT	National Institute of Genetics		
ESm library"	/dev_stage="3 days after infection with Meloidogyne	LOCATION	1111 Yata, Mishima, Shizuoka 411-8540, Japan		
/lab-host="XOLR"	/dev_stage="3 week old roots"	ORGANIZATION	Tei: 81-559-81-6856		
/db_xref="taxon:188702"	/dev_stage="3 days after infection with Meloidogyne	FEATURES	Fax: 81-559-81-6855		
/clone="PBNIR-4N2"	/dev_stage="3 week old roots"	SOURCE	Email: tshini@genes.nig.ac.jp.		
/clone_lib="PBNIR"	/dev_stage="3 days after infection with Meloidogyne	LOCATION/QUALIFIERS	Location/Qualifiers		
/clone_type="3 week old roots"	/dev_stage="3 days after infection with Meloidogyne	ORGANISM	1. 555 "Triticum aestivum"		
/dev_stage="3 days after infection with Meloidogyne	/dev_stage="3 days after infection with Meloidogyne	CULTIVAR	/organism="Chinese Spring"		
/note="Vector: plusescript SK- Site_1: EcorI; Site_2: XbaI; CDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-assist helper phage and propagated in XLOR cells."	/db_xref="taxon:4555"	VECTOR	/vector: Lambda Uni-ZAP XR, excised phagenid;		
BASE COUNT	185 a 117 c 151 g 186 t	CLONE	Site_1: XbaI; Plants were grown under hydroponic conditions at UC Davis, salt stressed for 12 hours, and for 7 days, then dissected and frozen (Akhnun in J Dvorak Lab). Total RNA was prepared from the two tissue, equal quantities of RNA were pooled from the two samples, PolyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give plasmid phagenids in the TJ Close lab at the University of California, Riverside (Akhnun, China).		
ORIGIN		LIBRARY	/clone lib="Y. Ogihara unpublished cDNA library wh_r"		
Query Match	4.9%	SEQUENCE	/tissue_type="Root"		
Best Local Similarity	46.3%	PRED.	/dev_stage="Feeke's scale 1"		
Matches	173;	MISMATCHES	/cultivar="Chinese Spring"		
Conservative	0;	INDELS	/note="Vector: Lambda Uni-ZAP XR, excised phagenid; Site_1: EcorI; Site_2: XbaI; Plants were grown under hydroponic conditions at UC Davis, salt stressed for 12 hours, then dissected and frozen (Akhnun in J Dvorak Lab). Total RNA was prepared from the two tissue, equal quantities of RNA were pooled from the two samples, PolyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give plasmid phagenids in the TJ Close lab at the University of California, Riverside (Akhnun, China). Choi, Close, Fenlon, Kianian, Otto, Simons, Zhang. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."		
Qy	184	MISMATCHES	/clone="wh_r"		
Db	13	0;	BASE COUNT	141 a 149 g 122 g 143 t	
Qy	244	0;	ORIGIN	Query Match 4.8%; Score 51.6; DB 13; Length 555; Best Local Similarity 47.5%; Pred. No. 0.0026; Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	73	0;	Qy	154 TACCCGACCCGAAAGAGTAAAGGGTTCTGGCGTTTAAG 213	
Qy	304	0;	Db	232 TACCTGTATCCCTGAAACGCCGACTTGCGTGTGCTGAAATCGGTGATGCTG 291	
Db	133	0;	Qy	214 GAGAAATTTAGTTGCTGGTACGGTGGACACTCATATACTACCTCAATAGCT 273	
Qy	364	0;	Db	292 ACTGAGCPACATCTGTTGGATGTGGTGTGATGAGCTATTGATGAGATGT 351	
Db	193	0;	Qy	244 AGAGAAAACCGTCTCGGGTACCTGGAAACCCACGGAAACCTT 483	
Qy	253	0;	Db	352 GTACTTGAAACGAGGTGACAAAATGTAGATGCCCTTCACCATCTGTT 411	

"Library" /note/-this sequence is part of a project of EST assemblies resulting from the application of public economics to seed-DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the Menger addressing of BACs in conjunction with the Maize "BACmap" project.

Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the ongoing addressing of BACs in conjunction with the Maize Mapping Project" BASE COUNT 467 a 398 c 431 g 461 t ORIGIN
 Query Match 4.4%; Score 47.2; DB 11; Length 1757; Best Local Similarity 44.8%; Pred. No. 0.058; Matches 181; Conservative 0; Mismatches 223; Indels 0; Gaps 0;
 Qy 154 TACCCAGACGCCGAAGCGAAAGAGTTAAAGGGGTCTTGC GGATTTTCGGCTTAAG 213
 Db 450 TACCCCTGATCCTGAAAGCCGCAATTGCGTCTGCCCTTGAGATCTGGACTTGA 509
 Qy 214 GAAAGAAATTTAGTCGTCGAACTCATATCTACCTCTCAATAGCT 273
 Db 510 TCTGATTACATACCTGCTGATGNGCGAAGATGACATATGAGAT 569
 Research 6: 791-806." BASE COUNT 126 a 213 c 198 g 125 t
 ORIGIN

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales
Chlamydomonadaceae; Chlamydomonas.

AUTHORS Crossman, A. **BASES** 1 to 662) **CHANGERS** C - W **DATES** 1940-1945 **HARRIS** E **HUNTER** G **LOCATIONS** 10000

TITLE: Analyses of the Chlamydomonas reinhardtii Genome: A Model. **AUTHORS:** P., McDermott, J.P., Shrager, J., Silflow, C., and Stern, D.; Grossman, A., Chang, C.-W., Davies, J.J., Harris, E., hauser, C., Lereboullet, A.

Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project; 1031

JOURNAL Unpublished (2001)

DuB
Duke University
Box 31000

Durham, NC 27708-1000

Tel: 919 613 8159
Fax: 919 613 8177
Email: chause@duke.edu.
FEATURES
source
Location/Qualifiers
1. f662
/organism="Chlamydomonas reinhardtii"

Search completed: April 15, 2003, 15:47:23

JOURNAL OF CLIMATE

